

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 11:19:49 ; Search time 12.48 Seconds  
(Without alignments)  
73.929 Million cell updates/sec

Title: US-09-252-828A-2

Sequence: 1 SWPVPQGFADICCCNKGDC.....TPSHSRQPHVMSQMSRSVS 41

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	97.1	424	1	US-08-453-472-7 Sequence 7, Appli
2	235	97.1	424	1	US-08-038-948-4 Sequence 4, Appli
3	235	97.1	424	1	US-08-453-952-7 Sequence 7, Appli
4	235	97.1	424	2	US-08-862-903-7 Sequence 7, Appli
5	235	97.1	424	2	US-08-484-158B-61 Sequence 61, Appli
6	226	93.4	223	2	US-08-484-993B-49 Sequence 49, Appli
7	226	93.4	223	2	US-08-484-158B-49 Sequence 49, Appli
8	226	93.4	223	2	US-08-484-596A-49 Sequence 49, Appli
9	226	93.4	223	2	US-08-480-150A-49 Sequence 49, Appli
10	226	93.4	223	3	US-08-458-731-49 Sequence 49, Appli
11	226	93.4	223	3	US-08-149-223A-49 Sequence 49, Appli
12	151	62.4	426	2	US-08-484-993B-12 Sequence 12, Appli
13	151	62.4	426	2	US-08-484-158B-12 Sequence 12, Appli
14	151	62.4	426	2	US-08-484-596A-12 Sequence 12, Appli
15	151	62.4	426	2	US-08-480-150A-12 Sequence 12, Appli
16	151	62.4	426	3	US-08-458-731-12 Sequence 12, Appli
17	151	62.4	426	3	US-08-149-223A-12 Sequence 12, Appli
18	136	56.2	424	2	US-08-484-993B-18 Sequence 18, Appli
19	136	56.2	424	2	US-08-484-158B-18 Sequence 18, Appli
20	136	56.2	424	2	US-08-484-596A-18 Sequence 18, Appli
21	136	56.2	424	2	US-08-480-150A-18 Sequence 18, Appli
22	136	56.2	424	3	US-08-458-731-18 Sequence 18, Appli
23	136	56.2	424	3	US-08-149-223A-18 Sequence 18, Appli
24	126	52.1	424	1	US-08-453-472-8 Sequence 8, Appli
25	126	52.1	424	1	US-08-038-948-2 Sequence 5, Appli
26	126	52.1	424	1	US-08-038-948-5 Sequence 5, Appli
27	126	52.1	424	1	US-08-453-952-8 Sequence 8, Appli

28	126	52.1	424	2	US-08-862-903-8 Sequence 8, Appli
29	119	49.2	421	2	US-08-484-993B-6 Sequence 6, Appli
30	119	49.2	421	2	US-08-484-158B-6 Sequence 6, Appli
31	119	49.2	421	2	US-08-484-596A-6 Sequence 6, Appli
32	119	49.2	421	2	US-08-480-150A-6 Sequence 6, Appli
33	119	49.2	421	3	US-08-458-731-6 Sequence 6, Appli
34	119	49.2	421	3	US-08-149-223A-6 Sequence 6, Appli
35	109	45.0	421	2	US-08-484-993B-24 Sequence 24, Appli
36	109	45.0	421	2	US-08-484-158B-24 Sequence 24, Appli
37	109	45.0	421	2	US-08-484-596A-24 Sequence 24, Appli
38	109	45.0	421	2	US-08-480-150A-24 Sequence 24, Appli
39	109	45.0	421	3	US-08-458-731-24 Sequence 24, Appli
40	109	45.0	421	3	US-08-149-223A-24 Sequence 24, Appli
41	100	41.3	415	2	US-08-484-158B-8 Sequence 8, Appli
42	100	41.3	415	2	US-08-484-993B-8 Sequence 8, Appli
43	100	41.3	415	2	US-08-484-596A-8 Sequence 8, Appli
44	100	41.3	415	2	US-08-480-150A-8 Sequence 8, Appli
45	100	41.3	415	3	US-08-458-731-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1  
US-08-453-472-7  
Sequence 7, Application US/08453472  
Patent No. 6626846  
GENERAL INFORMATION:  
APPLICANT: DEAN, JURRIEN  
TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
TITLE OF INVENTION: BASED ON ALLOIMUNIZATION WITH ZONA PELLUCIDA  
TITLE OF INVENTION: POLYPEPTIDES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,472  
FILING DATE: 30-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/038,948  
FILING DATE: 26-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/930,462  
FILING DATE: 20-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/364,379  
FILING DATE: 12-JUN-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4032 US3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein

# PROCEEDINGS

## ORGANIC FEATURES •

~~RESULT~~  
US-08-453



S-08-484-158B-61

Query Match 97.1%; Score 235; DB 2; Length 424;  
Best Local Similarity 95.1%; Pred. No. 7e-21;  
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 SWPVGADICCCCKGDCGTPSHSRPHVMSQMSRSVS 41  
b 308 SWPVGADICCCCKGDCGTPSHSRPHVMSQMSRSVS 348

RESULT 6

US-08-484-993B-49  
Sequence 49, Application US/08484993B

Patent No. 5837497

GENERAL INFORMATION:

APPLICANT: Harris Ph.D., Jeffrey D.

APPLICANT: Hsu, Kuang T.

APPLICANT: Podolski, Joseph S.

TITLE OF INVENTION: Materials and Methods for Immunocontraception

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,993B

FILING DATE: 09-NOV-1993

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/012,990

FILING DATE: 29-JAN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/973,341

FILING DATE: 09-NOV-1992

ATTORNEY/AGENT INFORMATION:

NAME: Clough, David W.

REGISTRATION NUMBER: 36,107

REFERENCE/DOCKET NUMBER: 31745

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6653

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 223 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-484-993B-49

Query Match 93.4%; Score 226; DB 2; Length 223;  
Best Local Similarity 90.2%; Pred. No. 4.5e-20;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 SWPVGADICCCCKGDCGTPSHSRPHVMSQMSRSVS 41  
b 107 SWPVGADICCCCKGDCGTPSHSRPHVMSQMSRSVS 147

RESULT 7

US-08-484-158B-49

Sequence 49, Application US/08484158B

Patent No. 5976545

GENERAL INFORMATION:

APPLICANT: Harris Ph.D., Jeffrey D.

APPLICANT: Hsu, Kuang T.

APPLICANT: Podolski, Joseph S.

TITLE OF INVENTION: Pharmaceutical Compositions for

TITLE OF INVENTION: Immunocontraception

NUMBER OF SEQUENCES: 61

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,158B

FILING DATE: 07-JUNE-95

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/149,223

FILING DATE: 09-NOV-93

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/012,990

FILING DATE: 29-JAN-93

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/973,341

FILING DATE: 09-NOV-92

ATTORNEY/AGENT INFORMATION:

NAME: Clough, David W.

REGISTRATION NUMBER: 36,107

REFERENCE/DOCKET NUMBER: 32794

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6653

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 223 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-484-158B-49

Query Match 93.4%; Score 226; DB 2; Length 223;  
Best Local Similarity 90.2%; Pred. No. 4.5e-20;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 SWPVGADICCCCKGDCGTPSHSRPHVMSQMSRSVS 41  
b 107 SWPVGADICCCCKGDCGTPSHSRPHVMSQMSRSVS 147

RESULT 8

US-08-484-596A-49

Sequence 49, Application US/08484596A

Patent No. 5981228

GENERAL INFORMATION:

APPLICANT: Harris Ph.D., Jeffrey D.

APPLICANT: Hsu, Kuang T.

APPLICANT: Podolski, Joseph S.

TITLE OF INVENTION: Materials and Methods for Immunocontraception

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive



NUMBER OF SEQUENCES: 01  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &



Thu Jan 10 11:47:25 2002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 426 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-480-150A-12

Query Match

Best local Similarity 62.4%; Score 151; DB 2; Length 426;

Matches 24; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Db 307 WYPVGSADICRCCKSGCLPGRSRRLSHLERGWRKSYS 346

Search completed: January 10, 2002, 11:20:40

Job time: 51 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OW protein - protein search, using sw model

Run on: January 10, 2002, 11:19:49 ; Search time 12.48 Seconds  
(without alignments)  
73.929 Million cell updates/s.

Title: US-09-252-828A-2

Perfect score: 242

Sequence: 1 SWPPVGPALICCCCKNGDC.....TPSHRRQPHYMSQMSRSVS 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents\_AA:\*

- 1: /cgcn2\_6/prodata/2/1aa/5A\_COMB.pep:\*
- 2: /cgcn2\_6/prodata/2/1aa/5B\_COMB.pep:\*
- 3: /cgcn2\_6/prodata/2/1aa/6A\_COMB.pep:\*
- 4: /cgcn2\_6/prodata/2/1aa/6B\_COMB.pep:\*
- 5: /cgcn2\_6/prodata/2/1aa/PCFUS\_COMB.pep:\*
- 6: /cgcn2\_6/prodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	97.1	424	1	US-08-453-472-7 Sequence 7, App
2	235	97.1	424	1	US-08-038-948-4 Sequence 4, App
3	235	97.1	424	1	US-08-453-952-7 Sequence 7, App
4	235	97.1	424	2	US-08-862-903-7 Sequence 7, App
5	235	97.1	424	2	US-08-484-158B-61 Sequence 61, App
6	235	97.1	424	2	US-08-484-993B-49 Sequence 49, App
7	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
8	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
9	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
10	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
11	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
12	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
13	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
14	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
15	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
16	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
17	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
18	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
19	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
20	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
21	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
22	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
23	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
24	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
25	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
26	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App
27	235	97.1	424	2	US-08-484-158B-49 Sequence 49, App

28	126	52.1	424	2	US-08-862-903-8	Sequence 8, App1
29	119	49.2	421	2	US-08-484-993B-6	Sequence 6, App1
30	119	49.2	421	2	US-08-484-158B-6	Sequence 6, App1
31	119	49.2	421	2	US-08-484-596A-6	Sequence 6, App1
32	119	49.2	421	2	US-08-480-150A-6	Sequence 6, App1
33	119	49.2	421	3	US-08-458-731-6	Sequence 6, App1
34	119	49.2	421	3	US-08-149-223A-6	Sequence 6, App1
35	109	45.0	421	2	US-08-484-993B-24	Sequence 24, App1
36	109	45.0	421	2	US-08-484-158B-24	Sequence 24, App1
37	109	45.0	421	2	US-08-484-596A-24	Sequence 24, App1
38	109	45.0	421	2	US-08-480-150A-24	Sequence 24, App1
39	109	45.0	421	3	US-08-458-731-24	Sequence 24, App1
40	109	45.0	421	3	US-08-149-223A-24	Sequence 24, App1
41	100	41.3	415	2	US-08-484-993B-8	Sequence 8, App1
42	100	41.3	415	2	US-08-484-158B-8	Sequence 8, App1
43	100	41.3	415	2	US-08-484-596A-8	Sequence 8, App1
44	100	41.3	415	2	US-08-480-150A-8	Sequence 8, App1
45	100	41.3	415	3	US-08-458-731-8	Sequence 8, App1

## ALIGNMENTS

RESULT 1  
US-08-453-472-7  
Sequence 7, Application US/08453472  
Patent No. 5626846  
GENERAL INFORMATION:  
APPLICANT: DEAN, JURRIEN  
TITLE OF INVENTION: CONTRACEPTIVE VACCINE  
TITLE OF INVENTION: BASED ON ALLOIMUNIZATION WITH ZONA PELLUCIDA  
TITLE OF INVENTION: POLYPEPTIDES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,472  
FILING DATE: 30-May-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/038,948  
FILING DATE: 26-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/364,379  
FILING DATE: 12-JUN-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4032 US3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein

```

ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: ZP3
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: human ZP3 protein
US-08-453-472-7

Query Match          97.1%; Score 235; DB 1; Length 424;
Best Local Similarity 95.1%; Pred. No. 7e-21;
Matches   39; Conservative    1; Mismatches    1; Indels    0; Gaps    0

OY      1 SWFVVGPDADICCCCKNKGDGTPSHSRQPHVMQWSRSYS 41
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db       308 SWFVEGPADICCCCKNGKDGCTPSHSRRQPHVMQWSRSAS 348

RESULT      2
US-08-038-948--4
Sequence 4, Application US/08038948
Patent No. 5641487
GENERAL INFORMATION:
APPLICANT: DEAN, JURRIEN
TITLE OF INVENTION: CONTRACEPTIVE VACCINE BASED ON
TITLE OF INVENTION: ALLOTIMIZATION WITH ZONA PELLUCIDA POLYPEPTIDES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 New York Avenue, N.W.
City: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/038,948
FILING DATE: 26-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,462
FILING DATE: 20-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: 99152/E-266-88/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 67146ZT CUSH
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

```

```

US-08-038-948-4
Query Match          97.1%; Score 235; DB 1; Length 424;
Best Local Similarity 95.1%; Pred. No. 7e-21;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 SWFPGVQPADICCCCNKGDCGTPSHSRQPHYMSQMSRSVS 41
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
DB 308 SWFPGVQPADICCCCNKGDCGTPSHSRQPHYMSQMSRSAS 348

1
US-08-453-952-7
Sequence 7, Application US/08453952
Patent No. 5672488
GENERAL INFORMATION:
APPLICANT: DEAN, JURRIEN
TITLE OF INVENTION: CONTRACEPTIVE VACCINE
TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,952
FILING DATE: 30-May-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/038,948
FILING DATE: 26-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,462
FILING DATE: 20-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4032 USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6649
TELEX: 421/92
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 424
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:

```

strain: 124  
 type: amino acid  
 strandness: single  
 topology: unknown  
 molecule type: protein  
 original source:  
 organism: human  
 strain:  
 individual isolate:  
 developmental stage:  
 haplotype:  
 tissue type:

```

1 TELEPHONE: 312/474-6653
2
3 TELEFAX: 312/474-0448
4
5 TEL#: 25-3856
6
7 INFORMATION FOR SEQ ID NO: 61:
8
9 SEQUENCE CHARACTERISTICS:
10
11 LENGTH: 424 amino acids
12
13 TYPE: amino acid
14
15 TOPOLOGY: linear
16
17 MOLECULE TYPE: protein
18
19 DESCRIPTION: /desc =
20
21 DESCRIPTION: "deduced amino acid sequence of human ZPC

```

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
332 South Wacker Drive  
Chicago, Illinois 60606



SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-458-731-49

Query Match 93.4%; Score 226; DB 3; Length 223;  
Best Local Similarity 90.2%; Pred. No. 4.5e-20;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 SWFVGGPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 41  
DB 107 SWFVGGPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 147

RESULT 11  
US-08-149-223A-49  
Sequence 49, Application US/08149223A  
Patent No. 6027727  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/149,223A  
FILING DATE: 09-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ. ID NO.: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-149-223A-49

Query Match 93.4%; Score 226; DB 3; Length 223;  
Best Local Similarity 90.2%; Pred. No. 4.5e-20;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWFVGGPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 41  
DB 107 SWFVGGPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 147

RESULT 12  
US-08-484-993B-12  
Sequence 12, Application US/08484993B  
Patent No. 5837497  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,993B  
FILING DATE: 09-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ. ID NO.: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 426 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-993B-12

Query Match 62.4%; Score 151; DB 2; Length 426;  
Best Local Similarity 60.0%; Pred. No. 8.3e-11;  
Matches 24; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

OY 2 WFPVGGPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 41  
DB 307 WFPVGGPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 346

RESULT 13  
US-08-484-158B-12  
Sequence 12, Application US/08484158B  
Patent No. 5976545  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Pharmaceutical Compositions for  
Immunocontraception  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

OY 1 SWFVGGPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 41  
DB 107 SWFVGGPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 147







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 11:20:44 ; Search time 10.18 seconds  
(without alignments)  
147.668 Million cell updates/sec

Title: US-09-252-828A-2

Perfect score: 242  
Sequence: 1 SWPPVGGPADICCCCKNGDC.....TPSHSRQPHVMSQWSRSVS 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

PCst-processing: Minimum Match 0%

Listing first 45 summaries

Database : SWISSProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	235	97.1	424	1	2P3A_HUMAN
2	229	94.6	424	1	2P3B_HUMAN
3	226	93.4	424	1	2P3A_MACRA
4	192	79.3	424	1	2P3A_CALSO
5	151	62.4	426	1	2P3_CANPA
6	143	59.1	422	1	2P3_MSAU
7	136	56.2	424	1	2P3_FELCA
8	126	52.1	424	1	2P3_MOUSE
9	119	49.2	421	1	2P3_PIG
10	109	45.0	421	1	2P3_BOVIN
11	100	41.3	415	1	2P3_RABIT
12	59	24.4	999	1	DSG3_HUMAN
13	56.5	23.3	819	1	STL_CHLMU
14	55.5	22.9	436	1	5H6_RAT
15	55	22.7	865	1	PMIL_HUMAN
16	55	22.7	867	1	PROM_MOUSE
17	54	22.3	202	1	P21_SOYBN
18	54	22.3	477	1	MA2_HUMAN
19	54	22.3	477	1	MA2_MOUSE
20	53.5	22.1	338	1	DEFB_HUMAN
21	52.5	21.7	447	1	ERL1_HORVU
22	52.5	21.7	447	1	ERL2_DAVCA
23	52.5	21.7	447	1	ERL2_HORVU
24	52.5	21.7	447	1	ERL2_HORVU
25	52.5	21.7	447	1	ERL2_HORVU
26	52.5	21.7	447	1	ERL2_HORVU
27	52.5	21.7	447	1	ERL2_HORVU
28	52.5	21.7	447	1	ERL2_HORVU
29	52.5	21.7	447	1	ERL2_HORVU
30	52.5	21.7	447	1	ERL2_HORVU
31	52.5	21.7	448	1	ERL2_HORVU
32	52.5	21.7	449	1	ERL2_HORVU
33	52.5	21.7	449	1	ERL2_HORVU

34	52.5	21.7	449	1	EF1A_MANES	049169 manihot esc
35	51.5	21.3	339	1	SYFA_CHLPN	094676 chlamydia p
36	51	21.1	65	1	TXW6_NAJNA	P29180 naja naja (
37	51	21.1	281	1	TRIZ_DROER	P34630 drosophila
38	51	21.1	1132	1	DNBI_HSV6U	P52338 human herpe
39	51	21.1	1132	1	DNBI_HSV62	P52538 human herpe
40	50.5	20.9	74	1	NXLD_ACAN	P34073 acanthopis
41	50.5	20.9	139	1	IGF_MYXGL	P22618 myxine glut
42	50.5	20.9	622	1	ACH4_CHICK	P09482 gallus gall
43	50	20.7	174	1	SODC_CAUCR	P20379 caulodacter
44	50	20.7	182	1	KRUC_SHEEP	P26372 ovis aries
45	50	20.7	269	1	ILIB_MOUSE	P10749 mus musculu

## ALIGNMENTS

RESULT	ID	STANDARD	PRT	424 AA.
1	2P3A_HUMAN			
AC	P21754			
DT	01-MAY-1991 (Rel. 18, Created)			
DT	01-MAY-1991 (Rel. 18, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	ZONA_PELLUCCIDA SPERM-BINDING PROTEIN 3A PRECURSOR (ZONA_PELLUCCIDA			
DE	GLYCOPROTEIN 2P3A) (ZONA_PELLUCCIDA PROTEIN C) (SPERM RECEPTOR) (ZP3).			
GN	2P3A OR ZP3.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90349545; PubMed=2385582;			
RA	Chamberlin M.E.; Dean J.;			
RT	"Human homolog of the mouse sperm receptor.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 87:6014-6018(1990).			
RN	[2]			
RP	SEQUENCE OF 329-424 FROM N.A.			
RC	TISSUE=Ovary;			
RX	MEDLINE=93122771; PubMed=1478648;			
RA	van Duin M., Polman J.E., Verkoelen C.C., Bunschoten H.,			
RT	Meyerink J.H., Olive W., Aitken R.J.;			
RT	"Cloning and characterization of the human sperm receptor ligand ZP3:			
RT	evidence for a second polymorphic allele with a different frequency.			
RT	In the Caucasian and Japanese populations.";			
RL	Genomics 14:1064-1070(1992).			
CC	- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR			
CC	SPERM-ADHESION TO THE ZONA_PELLUCCIDA, AND MAY CONTRIBUTE TO THE			
CC	SPECIES-SPECIFICITY OF THE INSEMINATION.			
CC	- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA_PELLUCCIDA, IN			
CC	WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.			
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR			
CC	MATRIX.			
CC	- PM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.			
CC	- SIMILARITY: CONTAINS 1 ZP DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
CC	EMBL: M60504; AAA61336.1; -			
CC	PIR: A36000; A36000.			
CC	MIM: 182889; -			
CC	InterPro: IPR001507; zona_pellucida.			
CC	Pfam: PF00100; zona_pellucida.1.			
CC	PRINTS: PR00023; ZPELUCCIDA.			
CC	SMART: SM00241; ZP.1.			
CC	PROSITE: PS00682; ZP_DOMAIN; 1.			

```

KW Glycoprotein; Signal; Sulfation; Sperm; Receptor;
KM Extracellular matrix; Multigene family.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 372 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3B.
FT DOMAIN 45 307 ZP.
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 372 AA; 41424 MW; AB78DB531BE7B4A CRC64;

Query Match      94.6%; Score 229; DB 1; Length 372;
Best Local Similarity 92.7%; Pred. No. 1,4e-20;
Matches 38; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db      1 SWFVPGPADTQCCKRGDCGTPSHSRQPHWNSOMRSYS 41
       |||||:|||||:|||||:|||||:|||||:|||||:|
       308 SWFPVEGPADICOCCKNGDCGTPSHSRQPHWNSOMSTAS 348

RESULT 3
ID      ZP3A_MACRA          STANDARD:      PRT;    424 AA.
ZP3A_MACRA
AC      P53785;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A PRECURSOR (ZONA PELLUCIDA
GN      GLYCOPROTEIN ZP3A) (ZONA PELLUCIDA PROTEIN C) (SPERM RECEPTOR) (ZP3).
OS      Macaca radiata (Bonnet monkey).
OC      Eukaryota; Metazoa; Chordata; Gracilata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OX      Ceropithecinae; Macaca.
NCBI_TaxID=9548;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96249321; PubMed=8848588;
RZ      Kolluri S.K., Kaul R., Banerjee K., Gupta S.K.;
RT      "Nucleotide sequence of cDNA encoding bonnet monkey (Macaca radiata)
RL      zona pellucida glycoprotein-zp3."
RZ      Reprod. Fertill. Dev. 7:1209-1212(1995).
CC      -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
CC      SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
CC      SPECIES-SPECIFICITY OF THE INSEMINATION.
CC      -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
CC      WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC      -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
CC      MATRIX.
CC      -1- PMM: SUPRATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.
CC      -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; X82639; CAA57961.1; -
DR      InterPro; IPR001507; zona_pellucida.
DR      Pfam; PF00100; zona_pellucida; 1.
DR      PRINTS; PR00023; ZPELLUCIDA.
DR      SMART; SM00241; zp; 1.
DR      ProSITE; PS00682; ZP_DOMAIN; 1.
KW      Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
KM      Extracellular matrix; Multigene family.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A.
FT DOMAIN 23 387 EXTRACELLULAR (POTENTIAL).

```

```

FT TRANSSEM 388 408 POTENTIAL.
FT DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 45 307 ZP.
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 424 AA; 47040 MW; 3B41C4CFA3792331 CRC64;

Query Match 93.4%; Score 226; DB 1; Length 424;
Best Local Similarity 90.2%; Pred. No. 3.6e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWFVPGPADICCCCKGDCGTPSHSRPHVMSQMSRSVS 41
D0 308 SWFVEGPADICCCCKGDCGTPSHSRPHVMSQMSRSAS 348

RESULT 4
ZP3A_CALSO
ID ZP3A_CALSO STANDARD: PRT: 424 AA.
AC P53786;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A PRECURSOR (ZONA PELLUCIDA GLYCOPROTEIN ZP3) (ZONA PELLUCIDA PROTEIN C) (SPERM RECEPTOR) (ZP3).
GN ZP3A OR ZP3.
OS Callithrix sp. (Marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callithrix.
OY NCBI_TaxId=9485;
RX [1]
RX SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RC MEDLINE=9436314; PubMed=8081814;
RA Thillai-Koothan P., van Duin M., Aitken R.J.;
RT "Cloning, sequencing and oocyte-specific expression of the marmoset sperm receptor protein, ZP3."
RL Zygote 1:93-101(1993).
CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE SPECIES-SPECIFICITY OF THE INSEMINATION.
CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR MATRIX.
CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S71825; AAB31866.1; -.
DR InterPro: IPR001507; zona.pellucida.
DR Pfam: PF00100; zona.pellucida; 1.
DR PRINTS: PR00023; ZPELUCIDA.
DR SMART: SM00241; ZP. 1.
DR PROSITE: PS00682; ZP_DOMAIN; 1.
KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane; Extracellular matrix; Multigene family.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A.
FT DOMAIN 23 387 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 388 408 POTENTIAL.
FT DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 307 ZP.

```

```

FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 424 AA; 46809 MW; 1DACBD03026C2739 CRC64;

Query Match 79.3%; Score 192; DB 1; Length 424;
Best Local Similarity 82.1%; Pred. No. 4.3e-16;
Matches 32; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 SWFVPGPADICCCCKGDCGTPSHSRPHVMSQMSRS 39
D0 308 SWFVEGPADICCCCKGDCGTPSHSRPHVMSLGSS 346

RESULT 5
ZP3_CANFA
ID ZP3_CANFA STANDARD: PRT: 426 AA.
AC P48831;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
GN ZP3 OR ZPC.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OY NCBI_TaxId=9615;
RX [1]
RX SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RC MEDLINE=95143578; PubMed=7841460;
RA Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C., Sacco A.G.;
RT "Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: the ZPA, ZPB and ZPC gene families."
RL DNA Seq. 4:361-393(1994).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Ovary;
RC Okazaki Y., Sugimoto M.;
RC Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY).
CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR MATRIX.
CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U05780; AAA74387.1; -.
DR EMBL: D45070; BAA08098.1; -.
DR InterPro: IPR001507; zona.pellucida.
DR Pfam: PF00100; zona.pellucida; 1.
DR PRINTS: PR00023; ZPELUCIDA.
DR SMART: SM00241; ZP. 1.
DR PROSITE: PS00682; ZP_DOMAIN; 1.
KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane; Extracellular matrix; Multigene family.
FT SIGNAL 1 22 POTENTIAL.

```

T	CHAIN	23	426	ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
T	DOMAIN	23	385	POTENTIAL.
T	TRANSMEM	386	406	EXTRACELLULAR (POTENTIAL).
T	DOMAIN	407	426	CYTOPLASMIC (POTENTIAL).
T	DOMAIN	43	305	NP.
T	CARBOHYD	123	123	N-LINKED (GLCNAC. . .) (POTENTIAL).
T	CARBOHYD	145	145	N-LINKED (GLCNAC. . .) (POTENTIAL).
T	CARBOHYD	244	244	N-LINKED (GLCNAC. . .) (POTENTIAL).
T	CONFLICT	227	227	L -> P (IN REF. 2).
T	CONFLICT	307	307	W -> S (IN REF. 2).
T	CONFLICT	343	343	K -> R (IN REF. 2).
SQ	SEQUENCE	426 AA:	47367 MW;	BES825A9ADCA172 CRC64;
Query Match	Similarity	62.4%;	Score 151;	DB L1; Length 426;
Best Local	Conservative	60.0%;	Freq. No. 3.6e-11;	
Matches	24;	5;	Mismatches	11;
I	Indels	0;	Gaps	0;
Oy	2 WEPYOGPADICOCCKNGDCCTPSHSRPHVMSQWSRSVS	41		
Dn	307 WYPEGASADICRCNCNKGCGLPGRSRRLSHLEGWRKRSVS	346		
RESULT	6			
ID	ZP3_MESAU	STANDARD;	PRT;	422 AA.
AC	P23491;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).			
GN	ZP3.			
OS	Mesocricetus auratus (Golden hamster).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;			
OC	Mesocricetus.			
OX	NCBI_TaxID=10036;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Ovary;			
RX	MEDLINE=91078540; PubMed=2257975;			
RA	Kinloch R.A., Ruiz-Seller B., Maassarman P.M.;			
RT	"Genomic organization and polypeptide primary structure of zona pellucida glycoprotein hzp3, the hamster sperm receptor.";			
RL	Dev. Biol. 142:414-421(1990).			
CC	- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE SPECIES-SPECIFICITY OF THE INSEMINATION.			
CC	- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.			
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR MATRIX.			
CC	- TISSUE SPECIFICITY: OOCYTES.			
CC	- DEVELOPMENTAL STAGE: GROWING OOCYTES.			
CC	- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.			
CC	- SIMILARITY: CONTAINS 1 ZP DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to llicense@isb-sib.ch).			
CC	-----			
DR	EMBL; M63629; AAA37079.1;			
DR	InterPro; IPRO01507; zona_pellucida.			
DR	Pfam; PF00100; zona_pellucida; 1.			
DR	PRINTS; PR00023; ZPELUCIDA.			
DR	SMART; SM00241; zp_1.			
DR	PROSITE; PS00687; ZP_DOMAIN; 1.			
KW	Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;			

KW		Extracellular matrix.	POTENTIAL.
FT	SIGMLL	1    22	ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
ET	CHAIN	23    422	EXTRACELLULAR (POTENTIAL).
FTT	DNAIEM	23    386	POTENTIAL.
FT	TRANSMEM	387    407	CYTOPLASMIC (POTENTIAL). .
FT	DOMAIN	408    422	.ZP.
FT	DOMAIN	45    306	PRO-RICH.
FT	DOMAIN	119    158	PRO-RICH.
FT	DOMAIN	208    257	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	146    146	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	271    271	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	302    302	N-LINKED (GLCNAC. .) (POTENTIAL).
SO	SEQUENCE	422 AA; 45827 MW;	DQF5BE/FHFE7E01 CRC64;
Query Match		59.1%:	Score 143; DB 1; Length 422;
Best Local Similarity		56.1%:	Pred. No. 3, 2e-10;
Matches	23; Conservative	7; Mismatches	11; Indels
			Gaps
OY	1 SWEPVPGPADICCCGNCKGDCGTPTSHSRDPHYMSOWSRSRV 41		
	: ::     :                 :		
DB	307 SMSPEVGDAIEVCGCCSSGDGCSSRRRYGAHVGSQMPKAS 347		

```

RESULT 7
ID ZP3_PELICA STANDARD; PRT; 424 AA.
DT P48832;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
GN ZP3 OR ZPC.
OS Fells silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Graniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Plissipedia; Fellidae; Fells.
OX NCBI_TaxID=9685;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RC MEDLINE=95133578; PubMed=7841460;
RX Harris J.D., Hildier D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
RT "Cloning and characterization of zona pellucida genes and cDNAs from
RT a variety of mammalian species: the ZPA, ZPB and ZPC gene families."
RT DNA Seq. 4:361-393(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RL Okazaki Y., Sugimoto M.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
CC SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY). IN
CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA. IN
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
CC MATRIX.
CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
CC EMBL: U05778; AAA74390.1; -
CC DR EMBL: D45068; BAA08096.1; -

```

```

CR InterPro: IPR001507; zona_pellucida.
CR Pfam: PF00100; zona_pellucida.1.
CR PRINTS: PR00023; ZPELUCIDA.
CR SMART: SM00241; ZP.1.
CR PROSITE: PS00682; ZP_DOMAIN; FALSE_NEG.
KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix; Multigene family.
FT SIGNAL 1 22
FT CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
FT DOMAIN 23 383 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 384 404 POTENTIAL.
FT DOMAIN 405 424 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 43 305 ZP.
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 72 72 G -> W (IN REF. 2).
FT CONFLICT 264 264 D -> Y (IN REF. 2).
SQ SEQUENCE 424 AA; 46853 MW; CFC62F35F9AFC0D CRC64;

Query Match 56.2%; Score 136; DB 1; Length 424;
Best Local Similarity 55.0%; Pred. No. 2.2e-09;
Matches 22; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 2 WFPVGPADICCCCKNGDCTPSSHRRPHVMSQMSRSVS 41
Db 307 WFPVGPADICCCCKNGSGCLOGRSWRLSHDRPMHKMS 346

RESULT 8
ZP3_MOUSE
ID ZP3_MOUSE STANDARD: PRT: 424 AA.
AC P10761;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
CN ZP3 OR ZP-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=88242926; PubMed=3378665;
RA Ringuette M.J., Chamberlin M.E., Baur A.W., Sobieski D.A., Dean J.;
RT "Molecular analysis of cDNA coding for ZP3, a sperm binding protein
RT of the mouse zona pellucida."
RL Dev. Biol. 127:287-293(1988).
RN [2]
RN REVISION TO 387.
RA Dean J.;
RN Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Liver;
RX MEDLINE=89240048; PubMed=2541416;
RA Kinloch R.A., Massarman P.M.;
RT "Nucleotide sequence of the gene encoding zona pellucida glycoprotein
RT ZP3 -- the mouse sperm receptor."
RN Nucleic Acids Res. 17:2861-2863(1989).
RN [4]
RN SEQUENCE FROM N.A.
RX MEDLINE=88320451; PubMed=2842770;
RA Kinloch R.A., Roller R.J., Fimiani C.M., Massarman D.A.;
RA Massarman P.M.;
RT "Primary structure of the mouse sperm receptor polypeptide determined
RT by genomic cloning."
RT Proc. Natl. Acad. Sci. U.S.A. 85:6409-6413(1988).
RN [5]
RN SEQUENCE OF 49-63; 197-204; 219-233 AND 261-275.

```

```

RC STRAIN=CD-1;
RX MEDLINE=93050795; PubMed=1330788;
RA Rosiere T.K., Massarman P.M.;
RT "Identification of a region of mouse zona pellucida glycoprotein mZP3
RT that possesses sperm receptor activity."
RL Dev. Biol. 154:309-317(1992).
CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
CC SPECIES-SPECIFICITY OF THE INSEMINATION.
CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
CC MATRIX.
CC -1- TISSUE SPECIFICITY: OOCYTES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE 2-WEEK GROWTH PHASE OF
CC OOGENESIS, PRIOR TO OVULATION.
CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M20026; AB18629.1; -.
DR EMBL: X14376; CA432550.1; -.
DR PIR: S04189; S04189.
DR PIR: A30334; A30334.
DR PIR: A31232; A31232.
DR MGD: MGI:99215; ZP3.
DR InterPro: IPR001507; zona_pellucida.
DR Pfam: PF00100; zona_pellucida.1.
DR PRINTS: PR00023; ZPELUCIDA.
DR SMART: SM00241; ZP.1.
DR PROSITE: PS00682; ZP_DOMAIN: 1.
KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix.
FT SIGNAL 1 22
FT CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
FT DOMAIN 23 387 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 388 408 POTENTIAL.
FT DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 308 ZP.
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 424 AA; 46303 MW; 9089903FBD268365 CRC64;

Query Match 52.1%; Score 126; DB 1; Length 424;
Best Local Similarity 53.7%; Pred. No. 3.6e-08;
Matches 22; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 SWFPVGPADICCCCKNGDCTPSSHRRPHVMSQMSRSVS 41
Db 309 SWLVEGDADICDCSHNCSSSSQFQIHGRPMKSLVS 349

RESULT 9
ZP3_PIG
ID ZP3_PIG STANDARD: PRT: 421 AA.
AC P42098;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3-BETA PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP3-BETA) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
GN ZP3B OR ZPC.

```

```
Sus scrofa (pig).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
NCBI_TaxID=9823;
(1)
SEQUENCE FROM N.A.
TISSUE-Ovary;
Yurewicz E.C., Hibler D., Fontenot G.K., Harris J.;
Submitted (JUL-1993) to the EMBL/Genbank/DBJ databases.
[2]
SEQUENCE FROM N.A.
TISSUE-Ovary;
Okazaki Y., Sugimoto M.;
Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.
-1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
SPERM-ADHESION TO THE ZONA PELLUCIDA. AND MAY CONTRIBUTE TO THE
SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY).
-1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA. IN
WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
MATRIX.
-1- TISSUE SPECIFICITY: OOCYTES.
-1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES
(BY SIMILARITY).
-1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC
EMBL; L22169; AAA31145.1;
DR EMBL; D45063; BAA08093.1;
DR GlycoSuiteDB; P42098;
DR InterPro; IPR001507; zona_pellucida.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00023; ZPELLUCIDA.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix.
FT SIGNAL 1 21
FT CHAIN 22 421
FT FT
FT DOMAIN 22 381
FT TRANSMEM 382 402
FT DOMAIN 403 421
FT DOMAIN 44 306
FT CARBOHYD 124 124
FT FT
FT CARBOHYD 146 146
FT FT
FT CARBOHYD 179 179
FT CARBOHYD 271 271
FT FT
FT CONFLICT 101 101
FT CONFLICT 107 107
FT CONFLICT 163 164
FT CONFLICT 404 404
FT SEQUENCE 421 AA; 46239 MW; DC97D760B985213A CRC64;
Query Match 49.2%; Score 119; DB 1; Length 421;
Best Local Similarity 70.4%; Pred. No. 2.4e-07;
Matches 19; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
```

```
RESULT 10
ZP3_BOVIN STANDARD; PRT; 421 AA.
ID ZP3_BOVIN
AC P48830;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP3B) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
DE ZP3 OR ZPC.
GN Bos taurus (Bovine)
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
(1)
SEQUENCE FROM N.A.
TISSUE-Ovary;
MEDLINE=95143578; PubMed=7841460;
Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
Sacco A.G.;
*Cloning and characterization of zona pellucida genes and cDNAs from
a variety of mammalian species: the ZPA, ZPB and ZPC gene families.*;
RA DNA Seq. 4:361-393(1994).
RT
RL
CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
CC SPERM-ADHESION TO THE ZONA PELLUCIDA. AND MAY CONTRIBUTE TO THE
CC SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY).
CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA. IN
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
CC MATRIX.
CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC
EMBL; U05775; AAA74385.1;
DR InterPro; IPR001507; zona_pellucida.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00023; ZPELLUCIDA.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix; Multigene family.
FT SIGNAL 1 22
FT CHAIN 23 421
FT DOMAIN 23 381
FT TRANSMEM 382 402
FT DOMAIN 403 421
FT DOMAIN 44 306
FT CARBOHYD 124 124
FT CARBOHYD 146 146
FT CARBOHYD 179 179
FT CARBOHYD 271 271
FT SEQUENCE 421 AA; 46545 MW; 905C4722B7B11DC CRC64;
Query Match 45.0%; Score 109; DB 1; Length 421;
Best Local Similarity 56.7%; Pred. No. 3.9e-06;
Matches 17; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
```

```

RESULT 11
223_RABIT STANDARD; PRT: 415 AA.
AC P48833;
DE 01-FEB-1996 (Rel. 33, Created)
DE 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-FEB-1996 (Rel. 33, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C)
DE (FRAGMENT).
GH ZP3 OR ZPC.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
ON NCBI_TaxID=9986;
RN [1]
RF SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=95143578; PubMed=7841460;
RA Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
RA Sacco A.G.;
RT Cloning and characterization of zona pellucida genes and cDNAs from
a variety of mammalian species: the ZPA, ZPB and ZPC gene families.";
RL DNA Seq. 4:361-393(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL: U05782; AAA74392.1;
DR InterPro: IPR001507; zona_pellucida.
DR Pfam: PF00100; zona_pellucida; 1.
DR SMART: SM00241; ZP_1.
DR PROSITE: PS00682; ZP_DOMAIN; 1.
KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix; Multigene family.
FT SIGNAL 1 18
FT CHAIN 19 415 POTENTIAL.
FT DOMAIN 19 378 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 379 399 POTENTIAL.
FT DOMAIN 400 415 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 41 301 ZP.
SQ SEQUENCE 415 AA; 44987 MW; 77396CF1BAA3P5CB CRC64;

Query Match 41.3%; Score 100; DB 1; Length 415;
Best local Similarity 42.6%; Pred. No. 4.6e-05;
Matches 20; Conservative 5; Mismatches 10; Indels 12; Gaps 2;

OY 1 SMPVQGPADICCCCKNGDC-----GTP-----SHSRQPHVMSQ 35
DB 302 SWAVEESADICECCGDCDLIAGSPMNHARSLSRHRVTEE 348

RESULT 12
DSG3_HUMAN STANDARD; PRT: 999 AA.
AC P32926;
DE 01-OCT-1993 (Rel. 27, Created)
DE 01-OCT-1993 (Rel. 27, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE DESMOGLEIN 3 PRECURSOR (130 KDA PEMPHIGUS VULGARIS ANTIGEN) (PVA).
EN DSG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92069753; PubMed=1720352;
RA Amagai M., Klaus-Kovtun V., Stanley J.R.;
RT Autoantibodies against a novel epithelial cadherin in pemphigus
RT vulgaris, a disease of cell adhesion.";
RL Cell 67:869-877(1991).
CC -1- FUNCTION: COMPONENT OF INTERCELLULAR DESMOsome JUNCTIONS.
CC INVOLVED IN THE INTERACTION OF PLAQUE PROTEINS AND INTERMEDIATE
CC FILAMENTS MEDIATING CELL-CELL ADHESION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EPIDERMIS, TONGUE, TONSIL, ESOPHAGUS AND
CC CARCINOMAS.
CC -1- DOMAIN: CALCIUM MAY BE BOUND BY THE CADHERIN-LIKE REPEATS
CC (POTENTIAL).
CC -1- DISEASE: PEMPHIGUS VULGARIS (PV) IS A POTENTIALLY LETHAL SKIN
CC DISEASE IN WHICH EPIDERMAL BLISTERS OCCUR AS THE RESULT OF THE
CC LOSS OF CELL-CELL ADHESION CAUSED BY THE ACTION OF AUTOANTIBODIES
CC AGAINST DSG3.
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL: M76482; AAA60230.1;
DR PIR: A41088; ITHU3.
DR HSSP: P09803; IEDH.
DR MIM: 169615;
DR InterPro: IPR002126; Cadherin.
DR Pfam: PF00028; cadherin; 4.
DR SMART: SM00112; CA; 4.
DR PROSITE: PS00232; CADHERIN_1; 3.
DR PROSITE: PS02658; CADHERIN_2; 4.
KW Cell adhesion; Signal; Transmembrane; Cytoskeleton; Glycoprotein;
KW Calcium-binding; Repeat.
FT SIGNAL 1 23
FT PROPEP 24 49
FT CHAIN 50 999
FT DOMAIN 50 615 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 616 640
FT DOMAIN 641 999 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 50 158 CADHERIN 1.
FT DOMAIN 159 268 CADHERIN 2.
FT DOMAIN 269 383 CADHERIN 3.
FT DOMAIN 386 499 CADHERIN 4.
FT REPEAT 910 935 DESMOGLEIN REPEAT 1.
FT REPEAT 936 966 DESMOGLEIN REPEAT 2.
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 999 AA; 107503 MW; 60479DD46AC219A1 CRC64;

Query Match 24.4%; Score 59; DB 1; Length 999;
Best local Similarity 39.3%; Pred. No. 7.9;
Matches 11; Conservative 4; Mismatches 7; Indels 6; Gaps 1;

OY 10 DICQCCNKGDGCT-----PSHSRRQPH 31
DB 585 EYCCDNRGIGIGTSYPTSPGTRGRPH 612

RESULT 13
SYL_CHLMU STANDARD; PRT: 819 AA.
AC Q9PK14;
DE 20-AUG-2001 (Rel. 40, Created)

```









---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 11:20:24 ; Search time 22.22 Seconds  
(without alignments)  
269,899 Million cell updates/sec

File: US-09-252-828a-2  
Perfect score: 242  
Sequence: 1 SWFPPVQGPADICCCNKGDC.....TPSHSRPQVHWSQMSRSVS 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues  
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_17:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp rodent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	217	89.7	210	4	Q12903	Q12903 homo sapien
2	125	51.7	424	11	P97708	P97708 rattus norv
3	115	47.5	418	11	Q9ER10	Q9ER10 microtus br
4	100.5	41.5	454	13	Q91675	Q91675 xenopus lae
5	100.5	41.5	460	13	Q91728	Q91728 xenopus lae
6	98.5	40.7	446	13	P79762	P79762 gallus gall
7	98.5	40.7	446	13	Q9PW68	Q9PW68 gallus gall
8	95	39.3	422	6	077685	077685 trichosurus
9	87	36.0	446	13	073670	073670 coturnix co
10	60.5	25.0	99	2	Q9PE89	Q9PE89 xyella fas
11	60.5	25.0	436	13	Q9DG37	Q9DG37 pimephales
12	59.5	24.6	934	2	005478	005478 streptomyce
13	58.5	24.2	559	13	Q98812	Q98812 brachydanio
14	57	23.6	1120	5	Q20778	Q20778 caenorhabdi
15	56	23.1	122	4	Q9P166	Q9P166 homo sapien
16	56	23.1	448	10	Q9FYV3	Q9FYV3 saccharum o
17	56	23.1	641	11	Q08463	Q08463 rattus norv
18	56	23.1	642	11	070421	070421 mus musculu
19	56	23.1	647	4	Q94815	Q94815 homo sapien

20	56	23.1	648	4	Q9UP38	Q9UP38 homo sapien
21	55.5	22.9	344	11	054788	054788 mus musculu
22	55.5	22.9	349	11	Q99N34	Q99N34 rattus norv
23	55.5	22.9	435	13	Q9PU06	Q9PU06 carassius a
24	55.5	22.9	447	10	Q9ZRP9	Q9ZRP9 malus domes
25	55	22.7	295	2	Q44588	Q44588 alcaligenes
26	55	22.7	591	11	Q9J149	Q9J149 rattus norv
27	54.5	22.5	340	12	Q9JF36	Q9JF36 vaccinia vl
28	54.5	22.5	424	13	Q92027	Q92027 cyprinus ca
29	54.5	22.5	759	13	Q919M5	Q919M5 xenopus lae
30	54.5	22.5	535	11	Q60473	Q60473 cavia porce
31	54	22.3	214	2	Q9RD56	Q9RD56 streptomyce
32	54	22.3	214	6	Q9RD56	Q9RD56 oryctolagus
33	54	22.3	315	6	Q97601	Q97601 oryctolagus
34	54	22.3	477	6	Q97600	Q97600 oryctolagus
35	54	22.3	537	4	Q9HBE2	Q9HBE2 homo sapien
36	54	22.3	544	1	Q48304	Q48304 haloterrax s
37	54	22.3	592	13	057328	057328 gallus gall
38	54	22.3	609	4	Q9HD61	Q9HD61 homo sapien
39	53.5	22.1	227	5	Q9VAG7	Q9VAG7 drosophila
40	53.5	22.1	270	4	Q9BRI6	Q9BRI6 homo sapien
41	53.5	22.1	338	4	Q60521	Q60521 homo sapien
42	53.5	22.1	351	1	Q9YA98	Q9YA98 aeropyrum p
43	53.5	22.1	379	10	Q9FYF9	Q9FYF9 arabidopsis
44	53.5	22.1	738	4	Q60327	Q60327 homo sapien
45	53.5	22.1	2168	5	Q19330	Q19330 caenorhabdi

## ALIGNMENTS

RESULT 1  
Q12903 PRELIMINARY; PRT; 210 AA.  
AC Q12903;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE POM-ZP3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=OVARY;  
RX MEDLINE=95309900; PubMed=7789967;  
RA Kiersztoz S., Osawa G.A., Liang L.F., Modi W.S., Dean J.;  
RT "POM-ZP3, a bipartite transcript derived from human ZP3 and a POM121  
homologue.";  
RL Genomics 25:354-359(1995).  
RT EMBL: U10099; AAA85788.1; -.  
DR InterPro: IPR001507; zona\_pellucida.  
DR Pfam: PF00100; zona\_pellucida; 1.  
SQ SEQUENCE 210 AA; 23196 MW; 6E9723EC1B77C8D0 CRC64;

Query Match 89.7%; Score 217; DB 4; Length 210;

Best Local Similarity 87.8%; Pred. No. 3,9e-22;

Matches 36; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SWFPPVQGPADICCCNKGDCGTPSHSRPQVHWSQMSRSVS 41  
Db 146 SWFPPVQGPADICCCNKGDCGTPSHSRPQVHWSQMSRSVS 186

RESULT 2  
P97708 PRELIMINARY; PRT; 424 AA.  
AC P97708; O55084;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR  
 DE (ZONA PELLUCIDA GLYCOPROTEIN ZP3) (ZONA PELLUCIDA GLYCOPROTEIN 3).  
 GN ZP3 OR ZP-3.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN NCBI\_TaxID=10116;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=WISTAR; TISSUE=OVARY;  
 RA Macduff P.E., Kerr L.F., Aitken R.J.;  
 RU J. Reprod. Fertil. Abstr. Ser. 18:86-86(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=SPRAGUE-DAWLEY; TISSUE=OVARY;  
 RA Akatsuka K., Yoshida-Komiyama H., Tulsiani D.P., Orgegin-Crist M.,  
 RA Hiroi M., Arai Y.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR  
 CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE  
 CC SPECIES-SPECIFICITY OF THE INSEMINATION.  
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN WHICH  
 CC ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR  
 CC MATRIX.  
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS A ZP DOMAIN, WHICH CURRENTLY HAS BEEN FOUND  
 CC IN ZP2, ZP3, GP2, TGFR-3 AND UROMODULIN.  
 CC EMBL: Y10823; CAA71787.1; -  
 CC EMBL: D78482; BAA24456.1; -  
 CC InterPro: IPR001507; zona\_pellucida.  
 DR Pfam: PR00100; zona\_pellucida; 1.  
 DR PRINTS: PR00023; ZPELLUCIDA.  
 DR SMART: SM00241; ZP; 1.  
 DR PROSITE: PS00682; ZP\_DOMAIN; FALSE NEG.  
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;  
 KW Extracellular matrix.  
 FT SIGNAL 1 16  
 FT CHAIN 17 424  
 FT DOMAIN 17 387  
 FT TRANSMEM 388 408  
 FT DOMAIN 409 424  
 FT DOMAIN 45 308  
 FT DOMAIN 329 334  
 FT CARBOHYD 146 146  
 FT CARBOHYD 273 273  
 FT CARBOHYD 304 304  
 FT CARBOHYD 327 327  
 FT CARBOHYD 330 330  
 FT CONFLICT 55 55  
 FT CONFLICT 112 112  
 FT CONFLICT 412 412  
 SQ SEQUENCE 424 AA; 45900 MW; 2AB42CB14DB701 CRC64;

Query Match 51.7%; Score 125; DB 11; Length 424;  
 Best Local Similarity 51.2%; Pred. No. 2.8e-09;  
 Matches 21; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

OY 1 SWFVPGPADICCCCKNGDCGTPSHSRROPHVMSQWSRSVS 41  
 DB 309 SWLPEVGDIDICCCSGNCSNSSSEFETHEPQWSTLV 349

RESULT 3  
 QSERIO PRELIMINARY; PRT; 418 AA.  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DT ZONA PELLUCIDA 3 GLYCOPROTEIN.

OS Microtus brandti.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Avicolinae;  
 OC Microtus.  
 RN NCBI\_TaxID=137787;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=OVARY;  
 RA Li H., Piao Y.S., Zhang Z.B., Zhu C., Hardy C., Hinds L.A.;  
 RT "Cloning and characterization of zona pellucida cDNA from Brandt's  
 RT vole."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF304487; ANG18455.1;  
 DR InterPro: IPR001507; zona\_pellucida.  
 DR Pfam: PF00100; zona\_pellucida; 1.  
 DR PRINTS: PR00023; ZPELLUCIDA.  
 DR SMART: SM00241; ZP; 1.  
 DR PROSITE: PS00682; ZP\_DOMAIN; UNKNOWN 1.  
 SQ SEQUENCE 418 AA; 45662 MW; 12C05062555F468F CRC64;

Query Match 47.5%; Score 115; DB 11; Length 418;  
 Best Local Similarity 44.7%; Pred. No. 6.5e-08;  
 Matches 17; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

OY 1 SWFVPGPADICCCCKNGDCGTPSHSRROPHVMSQWSR 38  
 DB 309 SWLPEVGDIDICCCCKNGDCSSRSYRPAHVAAPRNR 346

RESULT 4  
 QSERIO PRELIMINARY; PRT; 454 AA.  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE ZONA PELLUCIDA C GLYCOPROTEIN PRECURSOR.  
 GN XLZPC.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesodactylia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 RN NCBI\_TaxID=8335;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP Hedrick J.L., Yang J.C.;  
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U44952; AAB39079.1; -  
 DR InterPro: IPR001507; zona\_pellucida.  
 DR Pfam: PF00100; zona\_pellucida; 1.  
 DR SMART: SM00241; ZP; 1.  
 KW Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 454  
 SQ SEQUENCE 454 AA; 49589 MW; ADB9A3B978B5197 CRC64;

Query Match 41.5%; Score 100.5; DB 13; Length 454;  
 Best Local Similarity 46.3%; Pred. No. 6.7e-06;  
 Matches 19; Conservative 9; Mismatches 10; Indels 3; Gaps 3;

OY 1 SWFVQGPADICCCCKNGDCGTPSHSRROPHVMSQWSR 38  
 DB 330 SWSLPGGSMNICSCDITGNCVSPGOSRRLGPPYGSRRWD 370

RESULT 5  
 QSERIO PRELIMINARY; PRT; 460 AA.  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

```
D1 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE GP43.
O5 Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
OX Xenopodidae; Xenopus.
R3 NCBI_TaxID=8355;
R2 SEQUENCE FROM N.A.
R1 TISSUE-OVARIES FROM HCG-TREATED FEMALES;
R0 MEDLINE=98013448; PubMed=9352194;
RA Kubo H., Kawano T., Tsubuki S., Kawashima S., Katagiri C., Suzuki A.;
R3 "A major glycoprotein of Xenopus egg vitelline envelope, gp41, is a
R2 frog homolog of mammalian zp3."
R1 Dev Growth Differ. 39:405-417(1997).
R0 EMBL: D86568; BAA13117.1; -.
DK InterPro: IPR001507; zona_pellucida.
DK Pfam: PF00100; zona_pellucida; 1.
DK SMART: SM00241; ZP: 1.
SQ SEQUENCE 460 AA; 50084 MW; 3798F3E3394FDF34 CRC64;

Q2 1 SWPVOGPDICCCCKNGDC-GTPSHSR-QPHYM-SQMSR 38
DB 336 SWSPLOGSPNISCCTGNCVSPGOSRRLGPPYSGSRWQ 376

RESULT 6
P79762 PRELIMINARY; PRT; 446 AA.
AC P79762;
DE 01-MAY-1997 (TREMblrel. 03, Created)
DE 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, last annotation update)
D3 ZPC.
O5 Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae;
OX Gallus.
R3 NCBI_TaxID=9031;
R2 SEQUENCE FROM N.A.
R1 MEDLINE=99203469; PubMed=10103002;
RA Takeuchi Y., Nishimura K., Aoki N., Adachi T., Sato C., Kitajima K.,
R0 Matsuda T.;
R3 "A 42-kDa glycoprotein from chicken egg-envelope, an avian homolog of
R2 the ZPC family glycoproteins in mammalian zona pellucida. Its first
R1 identification, cDNA cloning and granulosa cell-specific expression.";
R0 Eur. J. Biochem. 260:736-742(1999).
D3 EMBL: D89097; BAA13760.2; -.
D2 InterPro: IPR001179; FKBP_PPIase.
D1 InterPro: IPR001507; zona_pellucida.
DK Pfam: PF00100; zona_pellucida; 1.
DK PRINTS: PR00023; ZPELUCIDA.
DK PROSITE: PS00453; FKBP_PPIASE_1; UNKNOWN_1.
DK SMART: SM00241; ZP: 1.
SQ SEQUENCE 446 AA; 47700 MW; DC5415232C22F03C CRC64;

Query Match 40.7%; Score 98.5; DB 13; Length 446;
Best Local Similarity 44.4%; Pred. No. 1.2e-05;
Matches 16; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

Q7 1 SWPVOGPDICCCCKNGDCGTPSHSRQPHYMSQW 36
DB 330 TWVPEGSRDVNCCTGNCCEPPALSRRL-NPMERW 364

RESULT 7
```

```
Q9PWF8 PRELIMINARY; PRT; 446 AA.
AC Q9PWF8;
DE 01-MAY-2000 (TREMblrel. 13, Created)
DE 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE ZONA PELLUCIDA C PROTEIN.
O5 Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae;
OX Gallus.
R3 NCBI_TaxID=9031;
R2 SEQUENCE FROM N.A.
R1 TISSUE-LIVER;
RA Kono Y., Matsuda T.;
R3 "Gallus gallus zona pellucida C protein gene.";
R2 Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
R0 EMBL: AB031033; BAA83418.1; -.
DR InterPro: IPR001179; FKBP_PPIase.
DR InterPro: IPR001507; zona_pellucida.
DK Pfam: PF00100; zona_pellucida; 1.
DK PRINTS: PR00023; ZPELUCIDA.
DK PROSITE: PS00453; FKBP_PPIASE_1; UNKNOWN_1.
DK SMART: SM00241; ZP: 1.
SQ SEQUENCE 446 AA; 47670 MW; DC5409532C22F03C CRC64;

QY 1 SWPVOGPDICCCCKNGDCGTPSHSRQPHYMSQW 36
DB 330 TWVPEGSRDVNCCTGNCCEPPALSRRL-NPMERW 364

RESULT 8
O77685 PRELIMINARY; PRT; 422 AA.
AC O77685;
DE 01-NOV-1998 (TREMblrel. 08, Created)
DE 01-NOV-1998 (TREMblrel. 08, last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, last annotation update)
D3 ZONA PELLUCIDA 3 PROTEIN.
GN ZP3.
O5 Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI_TaxID=9337;
R3 SEQUENCE FROM N.A.
R1 [1]
RA McCartney C.A., Mate K.E.;
R0 "Cloning and characterization of a zona pellucida 3 cDNA from a
R3 marsupial, the brush-tail possum Trichosurus vulpecula.";
R2 Zygote 0:0-0(1998).
R1 EMBL: AF079524; AAC28736.1; -.
DR InterPro: IPR001179; FKBP_PPIase.
DR InterPro: IPR001507; zona_pellucida.
DK Pfam: PF00100; zona_pellucida; 1.
DK PRINTS: PR00023; ZPELUCIDA.
DK PROSITE: PS00453; FKBP_PPIASE_1; UNKNOWN_1.
DK SMART: SM00241; ZP: 1.
SQ SEQUENCE 422 AA; 45714 MW; 58BA61A92F82612A CRC64;

Query Match 39.3%; Score 95; DB 6; Length 422;
Best Local Similarity 53.6%; Pred. No. 3.6e-05;
Matches 15; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 WFPVOGPDICCCCKNGDCGTPSHSRQ 29
DB 314 WLPVEGPRDICSCTGTCTISLSSRRK 341

RESULT 9
```

RA	Pelixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
RA	Quaglio R.E., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA	de Rosa V.B., Jr., de sa R.G., Santelli R.V., Sawasaki H.E.,
RA	da Silva A.C.F., da Silva A.M., da Silva F.R., Silva W.A., Jr.,
RA	da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA	de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshaho M.H.,
RA	Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
RA	Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT	"The genome sequence of the plant pathogen Xylella fastidiosa,"
RL	Nature 406:151-159(2000).
DR	EMBL; AE003949; AAF83949.1; -
KW	Hypothetical protein; Complete proteome.
SQ	SEQUENCE 99 AA; 11303 MW; 2CF720F47433FE81 CRC64;

  

Query Match	25.0%;	Score 60.5;	DB 2;	Length 99;
Best Local Similarity	45.8%;	Pred. No. 0.53;		
Matches 11; Conservative	3;	Mismatches 9;	Indels 1;	Gaps 1;

  

OY	12 CQCCKNGDCGTPSHSRRODPHWSQ 35
	:                 :
Dd	36 CRCARK-HCGLPHEHSRTFPNIPDQ 58

  

RESULT	11
ID	Q9DG37
	PRELIMINARY;
	PRT;
	436 AA.

Query Match		25.0%;	Score 60.5;	DB 2;	Length 99;
Best Local Similarity		45.8%;	Pred. No.0.53;		
Matches 11; Conservative		3;	Mismatches 9;	Indels 1;	Gaps 1;
OY	12 CQCCNKGCGRPSHRPVMVQS 35				
	:               :				
Db	36 CRCACK-HCGLPEHSRTFNPIDQ 58				
RESULT 11					
O9DG37	PRELIMINARY;	PRT;	436 AA.		
ID 09DG37					
AC 09DG37:					
DT 01-MAR-2001 (TREMBLrel. 16, Created)					
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)					
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)					
DE ZONA PELLUCIDA GLYCOPROTEIN 3 (FRAGMENT).					
GN zpz3.					
OS Pimephales promelas.					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;					
OC Cypriniformes; Cyprinidae; Leuciscinae; Pimephales.					
OX NCBI_TaxID=90988;					
XX [1]					
RN RP SEQUENCE FROM N.A.					
RA Morris C.A., Thomas-Jones R.E., Cryer J., Woodhead S., Kille P.;					
RT "Identification of the zp3 cDNA sequence from Pimephales promelas					
exposed to diethylstilbestrol.";					
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.					
DR EMBL; AF192407; AAC28398.1; -					
DR InterPro; IPR001507; zona_pellucida.					
DR Pfam; PF00100; zona_pellucida; 1.					
DR PRINTS; PR00023; ZPELLUCIDA.					
DR SMART; SMO0241; ZP; 1.					
FT NON_TER 1					
SQ SEQUENCE 436 AA; 47690 MW; 8C54DC948B641B6 CRC64;					
Query Match		25.0%;	Score 60.5;	DB 13;	Length 436;
Best Local Similarity		27.5%;	Pred. NO.1.9;		
Matches 11; Conservative		4;	Mismatches 24;	Indels 1;	Gaps 1;
OY	2 WFPVGPDADICQCCNKGCGRPSHRPVMVSQRSVS 41				
	:               :				
Db	385 WFADGNDQVCGCCD-STCGPDGERFAASPYGIEMEGRAS 423				
RESULT 12					
ID 005478	PRELIMINARY;	PRT;	934 AA.		
AC 005478:					
DT 01-JUL-1997 (TREMBLrel. 04, Created)					
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)					
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)					
DE LIPA TRANSCRIPTIONAL ACTIVATOR.					
GN LIPR.					
OS Streptomyces sp.					
OC Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae;					
OC					

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OK NCBI\_TaxID=1931;  
 RM (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=M1;  
 RX MEDLINE=93138418; PubMed=8422993;  
 RA Perez C., Juarez K., Garcia-Castells E., Soboron G.,  
 RA Servin-Gonzalez L.,  
 RC "Cloning, characterization, and expression in Streptomyces lividans 66  
 RC of an extracellular lipase-encoding gene from Streptomyces sp. M1.",  
 RL Gene 123:109-114(1993).  
 RM (2)  
 RE SEQUENCE FROM N.A.  
 RC STRAIN=M1;  
 RX Servin-Gonzalez L.,  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 DR EMBL: M86351; AAB51446.1; -;  
 DR InterPro: IPR000792; HTH\_LuxR.  
 DR Pfam: PF00196; Gere; 1.  
 DR SMART: SM00421; HTH\_LuxR; 1.  
 DR DNA-binding; Transcription regulation.  
 SQ SEQUENCE 934 AA; 95200 MW; 721770AAE7E8FA2 CRC64;

Query Match 24.6%; Score 59.5; DB 2; Length 934;  
 Best Local Similarity 31.4%; Pred. No. 5.2;  
 Matches 11; Conservative 6; Mismatches 17; Indels 1; Gaps 1;

OY 5 VOGPADICCC-CNKGDGCTPSHRSRPHYSOMSR 38  
 DB 641 VEGSGEACAHCDALAGAVPHGLAOPATLATYAR 675

RESULT 13  
 O9ES12 PRELIMINARY; PRT; 559 AA.  
 AC O98S12;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE FRIZZLED 7 PROTEIN (FRAGMENT).  
 EN F27.  
 SS Brachydanio rerio (zebrafish) (Zebra danio).  
 SC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 SC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 SC Cypriniformes; Cyprinidae; Rasbora; Danio.  
 NC NCBI\_TaxID=7955;  
 RX  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21184123; PubMed=11287199;  
 RX El-Messaoudi S., Renucci A.;  
 RT "Expression pattern of the frizzled 7 gene during zebrafish embryonic  
 RT development".  
 RL Mech. Dev. 102:231-234(2001).  
 RL EMBL: AJ301617; CAC37335.1; -;  
 RN NON\_TER 1  
 Q SEQUENCE 559 AA; 63719 MW; 25F243AA65F04D4B CRC64;

Query Match 24.2%; Score 58.5; DB 13; Length 559;  
 Best Local Similarity 43.3%; Pred. No. 4.5;  
 Matches 13; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

Y 3 PVOGPADICCCCNKGDGCTPSHRSRPHYV 32  
 B 147 FPVHGAGETICVGNSTSDAGSPT-SNPPYV 175

ESU\_JT 14  
 20778 PRELIMINARY; PRT; 1120 AA.  
 Q20778

AC 020778;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE SIMILAR TO TRIPLE HELICAL REGION OF COLLAGENS.  
 GN F54D8.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OK NCBI\_TaxID=6239;  
 RM (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,  
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans".  
 RL Nature 368:32-38(1994).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Bentley D.;  
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U12966; AAA20613.1; -;  
 DR InterPro: IPR00087; Collagen.  
 DR InterPro: IPR002486; COL\_cuticle\_N.  
 DR Pfam: PF01391; Collagen; 2.  
 DR Pfam: PF01484; COL\_cuticle\_N; 1.  
 SQ SEQUENCE 1120 AA; 119368 MW; 17764D8C4B88BCD CRC64;

Query Match 23.6%; Score 57; DB 5; Length 1120;  
 Best Local Similarity 36.7%; Pred. No. 13;  
 Matches 11; Conservative 1; Mismatches 10; Indels 8; Gaps 1;

OY 4 PVOGPADICCCCNKGDGCTPSH 25  
 DB 905 PSAGPGDKCNCNTENSCPAGPAGKPTGH 934

RESULT 15  
 O9P1G6 PRELIMINARY; PRT; 122 AA.  
 AC O9P1G6;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE PRO1787.  
 OS Homo sapiens (Human).  
 SC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 SC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NC NCBI\_TaxID=9606;  
 RX  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,  
 RA Liu M., He F.;  
 RT "Functional prediction of the coding sequences of 121 new genes  
 RT deduced by analysis of cDNA clones from human fetal liver.";

Submitted (DPC-1998) to the EMBL/GenBank/DBJ databases.  
EMBL: AF116668; AAF71088.1; B428ED1274CE81CD CRC64;  
SEQUENCE 122 AA; 13695 MW; 13695 MW;

Query Match 23.1%; Score 56; DB 4; Length 122;  
Best Local Similarity 36.6%; Pred. No. 2.6;  
Matches 15; Conservative 6; Mismatches 8; Indels 12; Gaps 2;  
6 QGPADI-----CQCCNKGDGCTPSHSRRORPHVMSQMSRSVS 41  
:|||||  
32 RGPADSSSHAFKLC-----FPHCRPRPPWVISRKGRAIN 65

Search completed: January 10, 2002, 11:23:28  
Job time: 184 sec



```
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
```

C:Accession: A56844  
 R:Kipeszczok, S.; Osawa, G.A.; Liang, L.; Modi, W.S.; Dean, J.  
 Genomics 25, 354-359, 1995  
 A:Title: POM-2P3, a bipartite transcript derived from human ZP3 and a POM121 homologue.  
 A:Reference number: A56844; MUID:95309900  
 A:Accession: A56844  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-210 <RIP>  
 A:Cross-references: GB:010099; NID:g607803; PIDN:AAA85788.1; PID:g607804

Query Match 89.7%; Score 217; DB 2; Length 210;  
 Best Local Similarity 87.8%; Pred. No. 9,7e-19;  
 Matches 36; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 SWPVGADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 41  
 DB 146 SWPVEGADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 186

RESULT 3  
 S70396  
 zona pellucida glycoprotein C - dog

C:Species: Canis lupus familiaris (dog)  
 C:Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 20-Aug-1999  
 C:Accession: S70396

R:Harris, J.D.; Hibler, D.W.; Fontenot, G.K.; Hsu, K.T.; Yurewicz, E.C.; Sacco, A.G.  
 DNA Seq. 4, 361-393, 1994

A:Title: Cloning and characterization of zona pellucida genes and cDNAs from a variety of  
 A:Reference number: S70396; MUID:95143578

A:Accession: S70396

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-426 <HAR>

A:Cross-references: EMBL:U05780; NID:g458276; PIDN:AAA74387.1; PID:g458277

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology  
 F:43-299/Domain: ZP domain homology <ZPH>

Query Match 62.4%; Score 151; DB 2; Length 426;  
 Best Local Similarity 60.0%; Pred. No. 1.2e-10;  
 Matches 24; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

OY 2 WFPVQPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 41  
 DB 307 WFPVGSADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 346

RESULT 4  
 A60503  
 sperm-binding glycoprotein ZP3 precursor - golden hamster

N:Alternate names: sperm receptor; zona pellucida glycoprotein ZP3

C:Species: Mesocricetus auratus (golden hamster)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A60503

R:Kinloch, R.A.; Ruiz-Seller, B.; Wassarman, P.M.  
 Dev. Biol. 142, 414-421, 1990

A:Title: Genomic organization and polypeptide primary structure of zona pellucida glycoprotein  
 A:Reference number: A60503; MUID:91078540

A:Accession: A60503

A:Molecule type: DNA

A:Residues: 1-422 <KIN>

A:Cross-references: GB:M63629

A>Note: the authors translated the codon CAA for residue 251 as Glu, and AGG for residue  
 C:Comment: This sulfated glycoprotein in the zona pellucida of the oocyte is a receptor

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

C:Keywords: glycoprotein; oocyte

F:45-300/Domain: ZP domain homology <ZPH>

Query Match 59.1%; Score 143; DB 1; Length 422;  
 Best Local Similarity 56.1%; Pred. No. 1e-09;

Matches 23; Conservative 7; Mismatches 11; Indels 0; Gaps 0;  
 OY 1 SWPVGADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 41  
 DB 307 SWPVEGADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 347

RESULT 5  
 S70399  
 zona pellucida glycoprotein C - cat

C:Species: Felis silvestris catus (domestic cat)

C:Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 20-Aug-1999

C:Accession: S70399

R:Harris, J.D.; Hibler, D.W.; Fontenot, G.K.; Hsu, K.T.; Yurewicz, E.C.; Sacco, A.G.  
 DNA Seq. 4, 361-393, 1994

A:Title: Cloning and characterization of zona pellucida genes and cDNAs from a variety of  
 A:Reference number: S70396; MUID:95143578

A:Accession: S70399

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-424 <HAR>

A:Cross-references: EMBL:U05778; NID:g458272; PIDN:AAA74390.1; PID:g458273

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology  
 F:43-299/Domain: ZP domain homology <ZPH>

Query Match 56.2%; Score 136; DB 2; Length 424;  
 Best Local Similarity 55.0%; Pred. No. 7.1e-09;  
 Matches 22; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 2 WFPVQPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 41  
 DB 307 WFPVEGADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 346

RESULT 6  
 A30334  
 sperm-binding glycoprotein ZP3 precursor - mouse

N:Alternate names: sperm receptor; zona pellucida glycoprotein ZP3; ZP3 glycoprotein

C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A30334; S04189; A31232; A48823

R:Ringue, M.J.; Chamberlin, M.E.; Baur, A.W.; Sobieski, D.A.; Dean, J.  
 Dev. Biol. 127, 287-295, 1988

A:Title: Molecular analysis of cDNA coding for ZP3, a sperm binding protein of the mouse  
 A:Reference number: A30334; MUID:88242926

A:Accession: A30334

A:Molecule type: DNA; mRNA

A:Residues: 1-424 <RIN>

A:Cross-references: GB:M20026; NID:g1663713; PIDN:AAI8629.1; PID:g1663714

R:Kinloch, R.A.; Wassarman, P.M.

Nucleic Acids Res. 17, 2861-2863, 1989

A:Title: Nucleotide sequence of the gene encoding zona pellucida glycoprotein ZP3 - t  
 A:Reference number: S04189; MUID:89240048

A:Accession: S04189

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-423 <KIN>

A:Cross-references: EMBL:X14376

R:Kinloch, R.A.; Roller, R.J.; Fimiani, C.M.; Wassarman, D.A.; Wassarman, P.M.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 6405-6413, 1988

A:Title: Primary structure of the mouse sperm receptor polypeptide determined by geno  
 A:Reference number: A31232; MUID:88320451

A:Accession: A31232

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-152, 'E', 154-252, 'E', 254-424 <KIN>

A:Cross-references: EMBL:J03851

R:Rosiere, T.K.; Wassarman, P.M.  
 Dev. Biol. 154, 309-317, 1992

A:Title: Identification of a region of mouse zona pellucida glycoprotein m2P3 that po  
 A:Reference number: A48823; MUID:93050795  
 A:Accession: A48823





A:Reference number: S45906  
 A:Accession: S78705  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-66 <ALT>  
 A:Cross-references: EMBL:Z35925; MIPS:YBR056w-a  
 A:Experimental source: strain S288C  
 C.Genetics:  
 A.Map position: 2R

Query Match 24.0%; Score 58; DB 2; Length 66;  
 Best Local Similarity 57.9%; Pred. No. 3.1;  
 Matches 11; Conservative 0; Mismatches 4; Indels 4; Gaps 1;  
 OX 6 QGPA----DICCCCNKGDC 20  
 ||| | | | | | |  
 DE 24 QGPPPRNDCCCCCNCGDC 42

Search completed: January 10, 2002, 11:21:00  
 Job time: 71 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: January 10, 2002, 11:19:49 ; Search time 11.74 Seconds  
(without alignments)  
118.044 Million cell updates/sec

Title: US-09-252-828A-2

Perfect score: 242  
Sequence: 1 SWFPVQGGADICCCCNKGDG.....TPSHSRKPHWSQMSRSVS 41

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 155098 seqs, 33800819 residues

Total number of hits satisfying chosen parameters: 155098

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/prodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/prodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	53.5	22.1	251	US-09-708-427-13464	Sequence 13464, A
2	53.5	22.1	296	US-09-708-427-13463	Sequence 13463, A
3	53.5	22.1	379	US-09-708-427-13462	Sequence 13462, A
4	53	21.9	165	US-09-708-427-80144	Sequence 80144, A
5	53	21.9	186	US-09-708-427-80143	Sequence 80143, A
6	53	21.9	210	US-09-708-427-80142	Sequence 80142, A
7	53	21.9	406	US-09-897-516-6739	Sequence 6739, Ap
8	52.5	21.7	174	US-09-620-394B-4805	Sequence 4805, Ap
9	52.5	21.7	227	US-09-620-394B-4804	Sequence 4804, Ap
10	52.5	21.7	275	US-09-620-394B-4803	Sequence 4803, Ap
11	52.5	21.7	348	US-09-620-394B-4932	Sequence 4932, Ap
12	52.5	21.7	348	US-09-620-394B-4932	Sequence 4932, Ap
13	52.5	21.7	348	US-09-708-427-22243	Sequence 22243, A
14	52.5	21.7	348	US-09-708-427-22247	Sequence 22247, A
15	52.5	21.7	399	US-09-708-427-58535	Sequence 58535, A
16	52.5	21.7	401	US-09-620-394B-3365	Sequence 3365, Ap
17	52.5	21.7	401	US-09-620-394B-4931	Sequence 4931, Ap
18	52.5	21.7	401	US-09-708-427-22239	Sequence 22239, A
19	52.5	21.7	401	US-09-708-427-22242	Sequence 22242, A
20	52.5	21.7	401	US-09-708-427-22246	Sequence 22246, A
21	52.5	21.7	447	US-09-708-427-58534	Sequence 58534, A
22	52.5	21.7	449	US-09-620-394B-3364	Sequence 3364, Ap
23	52.5	21.7	449	US-09-620-394B-4930	Sequence 4930, Ap
24	52.5	21.7	449	US-09-708-427-22238	Sequence 22238, A
25	52.5	21.7	449	US-09-708-427-22241	Sequence 22241, A
26	52.5	21.7	449	US-09-708-427-22245	Sequence 22245, A

## ALIGNMENTS

27	52.5	21.7	457	5	US-09-708-427-22237	Sequence 22237, A
28	52.5	21.7	484	5	US-09-708-427-58533	Sequence 58533, A
29	50.5	20.9	672	1	PCT-US01-27760-524	Sequence 524, App
30	50.5	20.9	672	1	PCT-US01-27760A-524	Sequence 524, App
31	50	20.7	170	5	US-09-708-427-70619	Sequence 70619, A
32	50	20.7	172	5	US-09-708-427-70619	Sequence 70619, A
33	50	20.7	204	5	US-09-708-427-82646	Sequence 82646, A
34	50	20.7	488	5	US-09-708-427-63568	Sequence 63568, A
35	50	20.7	535	5	US-09-708-427-85201	Sequence 85201, A
36	50	20.7	536	5	US-09-708-427-85200	Sequence 85200, A
37	50	20.7	565	5	US-09-708-427-85199	Sequence 85199, A
38	49.5	20.5	124	5	US-09-828-792-868	Sequence 868, App
39	49.5	20.5	154	5	US-09-828-792-987	Sequence 987, App
40	49.5	20.5	166	5	US-09-828-792-987	Sequence 987, App
41	49.5	20.5	453	5	US-09-989-723-73	Sequence 73, Appl
42	49.5	20.5	453	5	US-09-989-724-73	Sequence 73, Appl
43	49.5	20.5	453	5	US-09-989-730-73	Sequence 73, Appl
44	49.5	20.5	453	5	US-09-990-436-73	Sequence 73, Appl
45	49.5	20.5	453	5	US-09-997-641-73	Sequence 73, Appl

```

RESULT 1
US-09-708-427-13464
; Sequence 13464, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13464
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; LOCATION: 1..251
; NAME/KEY: misc.feature
; OTHER INFORMATION: Xaa is any amino acid
; LOCATION: 1..251
; OTHER INFORMATION: Ceres Seq. ID 1826393
US-09-708-427-13464

Query Match 22.1%; Score 53.5; DB 5; Length 251;
Best Local Similarity 32.3%; Pred. No. 15;
Matches 10; Conservative 5; Mismatches 13; Indels 3; Gaps 1;

QY 2 WFPVQGGADICCCCNKGDGPPSHSRQ 29
DB 158 WFAVRGGDCPGDGLRCSHAGCGPRTKRHE 188

RESULT 2
US-09-708-427-13463
; Sequence 13463, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13463
; LENGTH: 296

```

```

RESULT 6
US-09-708-427-80142
: Sequence 80142, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
: TITLE OF INVENTION: THEREBY
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 80142
: LENGTH: 210
: TYPE: PRT
: ORGANSIM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..210
: OTHER INFORMATION: Xaa is any amino acid

```



```

; NAME/KEY: misc_feature
; LOCATION: 1..210
; OTHER INFORMATION: Ceres Seq. ID 19599211
US-09-708-427-80142

```

Query Match	21.9%;	Score 53;	DB 5;	Length 210;
Best Local Similarity	30.3%;	Pred. NO. 15;		
Matches 10; Conservative	8;	Mismatches 15;	Indels 0;	Gaps 0;

```

ay 5 VQGPADLCQCKNKGGDCGTPSHSRQPHVMSQWS 37
    || :: || || || :: ||
Db 63 VQDAYVLCRYFHKNKNIIGPPSGNRYAFMEEEWA 95

```

```

RESULT      7
US-09-897-516-6739
; Sequence 6739, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 6739
; LENGTH: 406
; TYPE: PRF
; ORGANISM: Xenorhabdus sp.
; US-09-897-516-6739

```

Query Match	21.9%;	Score 53;	DB 5;	Length 406;
Best Local Similarity	28.9%;	Pred. NO. 25;		
Matches 13; Conservative	7;	Mismatches 13;	Indels 12;	Gaps 3

Q2 5 VQGPADICQCCNK-----GDCGTPSHSRQPHVMSQWSRSV 40  
 :|| :||| | :||| | ||: ::  
 D16 216 IQG-VRLCQCCATILPAQKHQCPRCHSKGHARR--HYSLQWTMAL 257

```

RESULT      8
US-09-620-394B-4805
: Sequence 4805, Application US/09620394B
:
: GENERAL INFORMATION:
: APPLICANT: ALEXANDROV, Nikolai
: APPLICANT: BROVER, Vyacheslav
: TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypeptides
: TITLE OF INVENTION: Thereby
: FILE REFERENCE: 2750-1067P
: CURRENT APPLICATION NUMBER: US/09/620,394B
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 9131
:
: SEQ ID NO 4805
:
: LENGTH: 174
:
: TYPE: PRT
:
: ORGANISM: Arabidopsis thaliana
:
: FEATURE:
: NAME/KEY: misc:feature
: LOCATION: 1..174
: OTHER INFORMATION: xaa is any amino acid
: NAME/KEY: misc:feature
: LOCATION: 1..174
: OTHER INFORMATION: Ceres Seq. ID 1393412

```

US-09-620-394B-4805

Query Match	21.78;	Score 52.5;	DB 5;	Length 174;
Best Local Similarity	58.88;	Pred. No. 15;		
Matches 10; Conservative	2;	Mismatches 4;	Indels 1;	Gaps 1;

```
QY 11 ICQCCNKGDGCTPSHSR 27
    ||||| ||:
Db 48 IC-CCNKMDATTPKYSK 63
```

```

RESULT      9
US-09-620-394B-4804
: Sequence 4804, Application US/09620394B
: GENERAL INFORMATION:
: APPLICANT: ALEXANDROV, Nikolai
: TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
: TITLE OF INVENTION: Thereby
: FILE REFERENCE: 2750-1067P
: CURRENT APPLICATION NUMBER: US/09/620,394B
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 9131
: SEQ ID NO 4804
: LENGTH: 227
: TYPE: PRP
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..227
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..227
: OTHER INFORMATION: Ceres Seq. ID 1393411
: US-09-620-394B-4804

```

Query Match	21.7%	Score 52.5;	DB 5;	Length 227;
Best Local Similarity	58.8%	Pred. No. 18;		
Matches 10;	Conservative 2;	Mismatches 4;	Indels 1;	Gaps 1;

```

QY      11 ICQCCNKGDCGTPSHSR 27
          ||||| | || :|:
Db      101 IC-CCNKMDATTPKYSK 116

```

```

RESULT 10
US-09-620-394B-4803
; Sequence 4803, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; TITLE OF INVENTION: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 4803
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..275
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..275
; OTHER INFORMATION: Ceres Seq. ID 1393410
; US-09-620-394B-4803

```

Query Match 21.7%; Score 52.5; DB 5; Length 275;  
Best Local Similarity 58.8%; Pred. No. 21;  
Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;  
OY 11 ICQCCNKGDGCTPSHR 27  
|| |||| | || :|  
Db 149 IC-CCNKMDATTPKYSK 164

RESULT 11  
US-09-620-394B-3366  
; Sequence 3366, Application US/09620394B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1067P  
; CURRENT APPLICATION NUMBER: US/09/620,394B  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 9131  
; SEQ ID NO 3366  
; LENGTH: 348  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..348  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..348  
; OTHER INFORMATION: Ceres Seq. ID 1385984  
US-09-620-394B-3366

Query Match 21.7%; Score 52.5; DB 5; Length 348;  
Best Local Similarity 58.8%; Pred. No. 26;  
Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

OY 11 ICQCCNKGDGCTPSHR 27  
|| |||| | || :|  
Db 48 IC-CCNKMDATTPKYSK 63

RESULT 12  
US-09-620-394B-4932  
; Sequence 4932, Application US/09620394B  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1067P  
; CURRENT APPLICATION NUMBER: US/09/620,394B  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 9131  
; SEQ ID NO 4932  
; LENGTH: 348  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..348  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..348  
; OTHER INFORMATION: Ceres Seq. ID 1393683  
US-09-620-394B-4932

Query Match 21.7%; Score 52.5; DB 5; Length 348;  
Best Local Similarity 58.8%; Pred. No. 26;  
Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

OY 11 ICQCCNKGDGCTPSHR 27  
|| |||| | || :|  
Db 48 IC-CCNKMDATTPKYSK 63

RESULT 13  
US-09-708-427-22243  
; Sequence 22243, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22243  
; LENGTH: 348  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..348  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..348  
; OTHER INFORMATION: Ceres Seq. ID 1840642  
US-09-708-427-22243

Query Match 21.7%; Score 52.5; DB 5; Length 348;  
Best Local Similarity 58.8%; Pred. No. 26;  
Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

OY 11 ICQCCNKGDGCTPSHR 27  
|| |||| | || :|  
Db 48 IC-CCNKMDATTPKYSK 63

RESULT 14  
US-09-708-427-22247  
; Sequence 22247, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22247  
; LENGTH: 348  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..348  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: 1..348  
; OTHER INFORMATION: Ceres Seq. ID 1840646  
US-09-708-427-22247

Query Match 21.7%; Score 52.5; DB 5; Length 348;  
Best Local Similarity 58.8%; Pred. No. 26;  
Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

OY 11 ICQCCNKGDGCTPSHR 27  
|| |||| | || :|

DJ 48 IC-CCNKMDDATTPKYSK 63

RESULT 15

US-09-708-427-58535

Sequence 58535, Application US/09708427

GENERAL INFORMATION:

APPLICANT: N. ALEXANDROV et al.

TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

FILE REFERENCE: 2750-1243P

CURRENT APPLICATION NUMBER: US/09/708,427

CURRENT FILING DATE: 2000-11-09

NUMBER OF SEQ ID NOS: 85364

SOFTWARE: PatentIn version 3.1

SEQ ID NO 58535

LENGTH: 399

TYPE: PRT

ORGANISM: Zea mays subsp. mays

FEATURE:

NAME/KEY: misc.feature

LOCATION: 1..399

OTHER INFORMATION: Xaa is any amino acid

NAME/KEY: misc.feature

LOCATION: 1..399

OTHER INFORMATION: Ceres Seq. ID 1942127

US-09-708-427-58535

Query Match 21.7%; Score 52.5; DB 5; Length 399;  
Best Local Similarity 58.8%; Pred. No. 28;

Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 11 ICOCCKNGDCCTPESHRSR 27

Db 101 IC-CCNKMDDATTPKYSK 116

Search completed: January 10, 2002, 11:21:19  
Job time: 90 sec

**THIS PAGE BLANK (USPTO)**



```
STREET: 20 South Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603
```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 19921109  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,497  
REFERENCE/DOCKET NUMBER: 31096  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/346-5750  
TELEFAX: 312/984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 426 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-973-341-12

Query Match 62.4%; Score 151; DB 3; Length 426;  
Best Local Similarity 60.0%; Pred.No. 3.2e-09;  
Matches 24; Conservative 5; Mismatches 11; Indels 0; Gaps 0.

DY 2 WPPVGGPADICCCCKKGDGCTPSHSRRQPHVMSRSRVS 41  
|::|||::||||| || | ||| :: | :|||  
Db 307 WYPEGSADICRCNKSGCLPGRSHSLTGRMKRSYS 346

RESULT 12  
US-08-259-580-12  
Sequence 12, Application US/08259580  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
TITLE OF INVENTION: DNA SEQUENCES ENCODING MAMMALIAN ZONA  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSER: Bicknell  
STREET: 20 South Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 19921109  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/012,990  
FILING DATE: 20 JAN 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,497

```

REFERENCE/DOCKET NUMBER: 31096
TELEPHONE: 312/346-5750
TELEFAX: 312/984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-259-580-12

Query Match                               62.4%; Score 151; DB 6; Length 426;
Best Local Similarity    60.0%; Pred. No. 3.2e-09;
Matches      24; Conservative       5; Mismatches   11; Indels     0; Gaps     0;

OY          2 WEPVQPADTCGCCNKGDCGTPSHSRQPHYMSQWSHSVS 41
             |:|:| ||||| |::| | | | | :|:||||
Db          307 WYPVGSGADICRCCNKGCGLPGRSRLSLHREGWRKSVS 346

RESULT 13
US-08-332-503-12
Sequence 12, Application US/08332503
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: DNA SEQUENCES ENCODING MAMMALIAN ZONA
NUMBER OF INVENTIONS: PELLUCIDA PROTEINS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Blacknell
STREET: 20 South Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/M5-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,503
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/973,341
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Botun, Michael F.
REGISTRATION NUMBER: 25,497
REFERENCE/DOCKET NUMBER: 31096
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/346-5750
TELEFAX: 312/984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-332-503-12

Query Match                               62.4%; Score 151; DB 7; Length 426;
Best Local Similarity    60.0%; Pred. No. 3.2e-09;
Matches      24; Conservative       5; Mismatches   11; Indels     0; Gaps     0;

```

STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,158A  
FILING DATE: 07-JUNE-95  
CLASSIFICATION: 514  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/149,223  
FILING DATE: 09-NOV-93  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-93  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-92  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 32794  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-158A-49

Query Match 93.4%; Score 226; DB 8; Length 223;  
Best Local Similarity 90.2%; Pred. No. 2.8e-18;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWFPVQGPADICCCNKGDCGTPSHSRPHVMSQMSRSVS 41  
Db 107 SWFPVQGPADICCCNKGDCGTPSHSRPHVMSQMSRSVS 147

RESULT 9  
US-09-252-828A-1  
Sequence 1, Application US/09252828A  
GENERAL INFORMATION:  
APPLICANT: DONG, KE-WEN  
APPLICANT: OEHNINGER, SERGIO  
APPLICANT: GIBBONS, WILLIAM  
TITLE OF INVENTION: RECOMBINANT, BIOLOGICALLY ACTIVE HUMAN ZONA PELLUCIDA  
FILE REFERENCE: 024754/0114  
CURRENT APPLICATION NUMBER: US/09/252,828A  
CURRENT FILING DATE: 1999-02-19  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 40  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-252-828A-1

Query Match 90.7%; Score 219.5; DB 16; Length 40;  
Best Local Similarity 95.1%; Pred. No. 3.5e-18;  
Matches 39; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SWFPVQGPADICCCNKGDCGTPSHSRPHVMSQMSRSVS 41  
Db 1 SWFPVQGPADICCCNKGDCGTPSHSRPHVMSQMSRSVS 40

RESULT 10  
PCT-US93-10851-12  
Sequence 12, Application PC/TUS9310851  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Materials and Methods for Immunococtraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/10851  
FILING DATE: 09-NOV-1993  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 426 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-10851-12

Query Match 62.4%; Score 151; DB 1; Length 426;  
Best Local Similarity 60.0%; Pred. No. 3.2e-09;  
Matches 24; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 2 WFPVQGPADICCCNKGDCGTPSHSRPHVMSQMSRSVS 41  
Db 307 WFPVQGPADICCCNKGDCGTPSHSRPHVMSQMSRSVS 346

RESULT 11  
US-07-973-341-12  
Sequence 12, Application US/07973341  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: DNA SEQUENCES ENCODING MAMMALIAN ZONA  
TITLE OF INVENTION: PELLUCIDA PROTEINS  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell

LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
CT-US93-10851-49

Query Match 93.4% Score 226; DB 1; Length 223;  
Best Local Similarity 90.2%; Pred. No. 2,8e-18;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 SWPVGADICCCCKNGDGTSPHSRQPHVMSQMSRSVS 41  
|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 107 SWPVEGADICCCCKSKDGTSPHSRQPHVVSQMSRSAS 147

RESULT 6  
US-08-480-444B-49  
Sequence 49, Application US/08480444B  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immuncontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,444B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,223  
FILING DATE: 09-NOV-1993  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-480-444B-49

Query Match 93.4% Score 226; DB 8; Length 223;  
Best Local Similarity 90.2%; Pred. No. 2,8e-18;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
1 SWPVGADICCCCKNGDGTSPHSRQPHVMSQMSRSVS 41  
|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 107 SWPVEGADICCCCKSKDGTSPHSRQPHVVSQMSRSAS 147

RESULT 7  
US-08-480-673B-49  
Sequence 49, Application US/08480673B  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immuncontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,673B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,223  
FILING DATE: 09-NOV-1993  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-480-673B-49

Query Match 93.4% Score 226; DB 8; Length 223;  
Best Local Similarity 90.2%; Pred. No. 2,8e-18;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 SWPVGADICCCCKNGDGTSPHSRQPHVMSQMSRSVS 41  
|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 107 SWPVEGADICCCCKSKDGTSPHSRQPHVVSQMSRSAS 147

RESULT 8  
US-08-484-158A-49  
Sequence 49, Application US/08484158A  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Pharmaceutical Compositions for  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun



RESULT 2  
US-09-252-828A-2  
Sequence 2, Application US/09252828A  
GENERAL INFORMATION:  
APPLICANT: DONG, KE-WEN  
APPLICANT: OEHNINGER, SERGIO  
APPLICANT: GIBBONS, WILLIAM  
TITLE OF INVENTION: RECOMBINANT, BIOLOGICALLY ACTIVE HUMAN ZONA PELLUCIDA  
TITLE OF INVENTION: PROTEIN 3 (hZP3) TO TEST MALE FERTILITY  
FILE REFERENCE: 024754/0114  
CURRENT APPLICATION NUMBER: US/09/252,828A  
CURRENT FILING DATE: 1999-02-19  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 41  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-252-828A-2

Query Match 100.0%; Score 242; DB 16; Length 41;  
Best Local Similarity 100.0%; Pred. No. 8,2e-21;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWFVVGPDADICCCCKNGDGGTSPSHSRPHVMSQMSRSVS 41  
|||||  
DB 1 SWFVVGPDADICCCCKNGDGGTSPSHSRPHVMSQMSRSVS 41

RESULT 3  
US-08-030-177-2  
Sequence 2, Application US/08030177  
GENERAL INFORMATION:  
APPLICANT: van DUIN, Marcel  
TITLE OF INVENTION: Human Zona pellucida protein ZP3  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akzo Pharma  
STREET: 1330-A Piccard Drive  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850-4377  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/030,177  
FILING DATE: 26-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 90,202,287.0  
FILING DATE: 27-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Blackstone, William M.  
REGISTRATION NUMBER: 29,772  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-030-177-2

Query Match 94.6%; Score 229; DB 4; Length 372;  
Best Local Similarity 92.7%; Pred. No. 2e-18;  
Matches 38; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWFVVGPDADICCCCKNGDGGTSPSHSRPHVMSQMSRSVS 41

|||||  
DB 308 SWFVVGPDADICCCCKNGDGGTSPSHSRPHVMSQMSRSVS 348

RESULT 4  
US-09-720-282-2  
Sequence 2, Application US/09720282  
GENERAL INFORMATION:  
APPLICANT: Herr, John  
APPLICANT: Coonrod, Scott  
TITLE OF INVENTION: EGG-SURFACE PROTEINS AND METHODS OF THEIR USE FOR MODULATING  
FERTILITY  
FILE REFERENCE: 9426-004-999  
CURRENT APPLICATION NUMBER: US/09/720,282  
CURRENT FILING DATE: 2001-01-18  
PRIOR APPLICATION NUMBER: PCT/US99/13858  
PRIOR FILING DATE: 1999-06-18  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 372  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-720-282-2

Query Match 94.6%; Score 229; DB 21; Length 372;  
Best Local Similarity 92.7%; Pred. No. 2e-18;  
Matches 38; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWFVVGPDADICCCCKNGDGGTSPSHSRPHVMSQMSRSVS 41  
|||||  
DB 308 SWFVVGPDADICCCCKNGDGGTSPSHSRPHVMSQMSRSVS 348

RESULT 5  
PCT-US93-10851-49  
Sequence 49, Application PC/TUS9310851  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/10851  
FILING DATE: 09-NOV-1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 28-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

DM protein - protein search, using sw model

Run on: January 10, 2002, 11:19:49 ; Search time 93.69 seconds  
(without alignments)  
121.506 Million cell updates/sec

File: US-09-252-828A-2

Perfect score: 242  
Sequence: 1 SWFPVGPADICCCCKNGDC.....TPSHSRQPHVMSQMSRSVS 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 27657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/ptodata/2/paa/PCRTUS\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	242	100.0	41	16	US-09-252-828A-1
2	242	100.0	41	16	US-09-252-828A-2
3	229	94.6	372	4	US-08-030-177-2
4	229	94.6	372	21	US-09-720-282-2
5	226	93.4	223	1	PCT-US93-10851-49
6	226	93.4	223	8	US-08-480-444B-49
7	226	93.4	223	8	US-08-480-673B-49
8	226	93.4	223	8	US-08-484-158A-49
9	219.5	90.7	40	16	US-09-252-828A-1

10	151	62.4	426	1	PCT-US93-10851-12	Sequence 12, Appl
11	151	62.4	426	3	US-07-973-341-12	Sequence 12, Appl
12	151	62.4	426	6	US-08-259-580-12	Sequence 12, Appl
13	151	62.4	426	7	US-08-332-503-12	Sequence 12, Appl
14	151	62.4	426	8	US-08-480-444B-12	Sequence 12, Appl
15	151	62.4	426	8	US-08-480-673B-12	Sequence 12, Appl
16	151	62.4	426	8	US-08-484-158A-12	Sequence 12, Appl
17	136	56.2	424	1	PCT-US93-10851-18	Sequence 18, Appl
18	136	56.2	424	1	PCT-US93-10851-18	Sequence 18, Appl
19	136	56.2	424	6	US-08-259-580-18	Sequence 18, Appl
20	136	56.2	424	6	US-08-332-503-18	Sequence 18, Appl
21	136	56.2	424	8	US-08-480-444B-18	Sequence 18, Appl
22	136	56.2	424	8	US-08-480-673B-18	Sequence 18, Appl
23	136	56.2	424	8	US-08-484-158A-18	Sequence 18, Appl
24	129	53.3	41	16	US-09-252-828A-2	Sequence 2, Appl
25	129	53.3	41	16	US-09-252-828A-3	Sequence 3, Appl
26	119	49.2	421	3	PCT-US93-10851-6	Sequence 6, Appl
27	119	49.2	421	3	US-07-973-341-6	Sequence 6, Appl
28	119	49.2	421	6	US-08-259-580-6	Sequence 6, Appl
29	119	49.2	421	7	US-08-332-503-6	Sequence 6, Appl
30	119	49.2	421	8	US-08-480-444B-6	Sequence 6, Appl
31	119	49.2	421	8	US-08-484-158A-6	Sequence 6, Appl
32	119	49.2	421	8	US-08-480-673B-6	Sequence 6, Appl
33	109	45.0	421	1	PCT-US93-10851-24	Sequence 24, Appl
34	109	45.0	421	3	US-07-973-341-24	Sequence 24, Appl
35	109	45.0	421	6	US-08-259-580-24	Sequence 24, Appl
36	109	45.0	421	7	US-08-332-503-24	Sequence 24, Appl
37	109	45.0	421	8	US-08-480-444B-24	Sequence 24, Appl
38	109	45.0	421	8	US-08-484-158A-24	Sequence 24, Appl
39	109	45.0	421	8	US-08-480-673B-24	Sequence 24, Appl
40	106	43.8	18	4	US-08-030-177-7	Sequence 8, Appl
41	100	41.3	415	1	PCT-US93-10851-8	Sequence 8, Appl
42	100	41.3	415	3	US-07-973-341-8	Sequence 8, Appl
43	100	41.3	415	6	US-08-259-580-8	Sequence 8, Appl
44	100	41.3	415	7	US-08-332-503-8	Sequence 8, Appl
45	100	41.3	415	8	US-08-480-444B-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1  
US-09-252-828-1  
; Sequence 1, Application US/09252828  
; GENERAL INFORMATION:  
; APPLICANT: DONG, Ke-Wen  
; APPLICANT: GIBBONS, William E.  
; TITLE OF INVENTION: RECOMBINANT, BIOLOGICALLY ACTIVE HUMAN ZONA PELLUCIDA  
; FILE REFERENCE: 024754/0114  
; CURRENT APPLICATION NUMBER: US/09/252,828  
; EARLIER FILING DATE: 1999-02-19  
; EARLIER APPLICATION NUMBER: US 60/075,079  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-252-828-1

Query Match 100.0%; Score 242; DB 16; Length 41;  
Best Local Similarity 100.0%; Pred. No. 8.2e-21;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SWFPVGPADICCCCKNGDCGTPSHSRQPHVMSQMSRSVS 41  
DB 1 SWFPVGPADICCCCKNGDCGTPSHSRQPHVMSQMSRSVS 41

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OMI protein - protein search, using sw model

Run on: January 10, 2002, 11:19:49 ; Search time 23.82 Seconds  
(without alignments)  
127.498 Million cell updates/sec

Title: US-09-252-828A-2

Perfect score: 242

Sequence: 1 SWFPVGPADICCCCKNGDC.....TPSHSRPHVMSQMSRSVS 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq.1101.\*

1:	/SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2:	/SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3:	/SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4:	/SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5:	/SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6:	/SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7:	/SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8:	/SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9:	/SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10:	/SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11:	/SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12:	/SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13:	/SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14:	/SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15:	/SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16:	/SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17:	/SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18:	/SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19:	/SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20:	/SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21:	/SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22:	/SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	Score	Query Match	Length	DB	ID	Description
1	242	100.0	41	20	AAV30214	Amino acids 308 to
2	235	97.1	424	21	AAV52988	Human zona pelluci
3	229	94.6	372	21	AAV52987	Human zona pelluci
4	226	93.4	223	15	AAV42483	Cynomolgus monkey
5	226	93.4	223	20	AAV42483	Cynomolgus monkey
6	226	93.4	223	20	AAV42483	Cynomolgus monkey
7	226	93.4	223	21	AAV42483	Cynomolgus monkey
8	226	93.4	223	21	AAV42483	Cynomolgus monkey
9	226	93.4	223	21	AAV42483	Cynomolgus monkey
10	226	93.4	223	21	AAV42483	Cynomolgus monkey
11	226	93.4	223	21	AAV42483	Cynomolgus monkey

12	192	79.3	424	15	AAV53498
13	151	62.4	426	15	AAV51199
14	151	62.4	426	20	AAV42472
15	151	62.4	426	20	AAV42472
16	151	62.4	426	21	AAV42472
17	151	62.4	426	21	AAV42472
18	151	62.4	426	21	AAV42472
19	151	62.4	426	21	AAV42472
20	140	57.9	426	15	AAV46068
21	137	56.6	424	15	AAV460165
22	137	56.6	424	15	AAV47198
23	136	56.2	424	15	AAV5202
24	136	56.2	424	20	AAV42475
25	136	56.2	424	20	AAV42475
26	136	56.2	424	21	AAV42475
27	136	56.2	424	21	AAV42475
28	136	56.2	424	21	AAV42475
29	136	56.2	424	21	AAV42475
30	126	52.1	424	11	AAV07058
31	119	49.2	258	14	AAV41004
32	119	49.2	258	15	AAV46951
33	119	49.2	420	15	AAV46950
34	119	49.2	421	15	AAV46950
35	119	49.2	421	20	AAV42469
36	119	49.2	421	20	AAV42469
37	119	49.2	421	21	AAV42469
38	119	49.2	421	21	AAV42469
39	119	49.2	421	21	AAV42469
40	119	49.2	421	21	AAV42469
41	109	45.0	421	15	AAV5205
42	109	45.0	421	20	AAV42478
43	109	45.0	421	20	AAV42478
44	109	45.0	421	21	AAV42478
45	109	45.0	421	21	AAV42478

## ALIGNMENTS

RESULT 1

AAV30214  
ID AAV30214 standard; peptide: 41 AA.

AAV30214:  
01-NOV-1999 (first entry)

Amino acids 308 to 348 of human zona pellucida protein 3 (hZP3).

Human: zona pellucida protein 3; hZP3; glycoprotein; human spermatozoa; acrosome reaction; fertility; sperm motility; fertilization.

Homo sapiens.  
MO9942581-A1.  
26-AUG-1999.

19-FEB-1999; 99WO-US03273.  
19-FEB-1998; 98US-0075079.

(EVIR-) EASTERN VIRGINIA MEDICAL SCHOOL.  
Dong K, Gibbons WE, Oehninger S;  
WPI: 1999-527476/44.

New recombinant human zona pellucida protein 3 - used to develop products for diagnostic and therapeutic use related to human fertility

Claim 7; Fig 1; 35pp; English.



DE 308 swfpevgpadicgccknkgdcgtpshrrqphvmsqwsstas 348

## RESULT 4

AA066552 standard; Protein; 223 AA.

AA066552;

03-FEB-1995 (first entry)

Cynomolgus Monkey zona pellucida ZPC protein.

Cynomolgus Monkey; zona pellucida; ZPC; immunocontraception.

Macaca cynomolgus.

WO9411019-A.

26-MAY-1994.

06-NOV-1993; 93WO-US10851.

09-NOV-1992; 92US-0973341.

29-JAN-1993; 93US-0012990.

(ZONA-) ZONAGEN INC.

Harris JD, Hsu KT, Podolski JS;

WPI: 1994-183156/22.

N-PSDB; AA079682.

Use of zona pellucida proteins and antibodies - for inducing

reproducible transient infertility or permanent sterility in

female mammals

Example 12: Page 132-133; 154pp; English.

Cynomolgus monkey cDNA libraries were constructed in lambda gt10

using mRNA isolated from a set of ovaries collected from monkeys

aged 1.5 and 2 years and a second set from monkeys aged 3, 4 and 14

years of age. The libraries were screened with probes encoding

porcine ZPA, ZPB and ZPC proteins. Positive clones were analysed

further by Southern hybridisation using the porcine probes. Clones

encoding cynomolgus monkey ZPA, ZPB and ZPC proteins were

identified (see AA079680-079682). AA079682 is the insert from the

largest partial ZPC clone which contains just over 50% of the

C-terminal portion of the full-length sequence and contains an ORF

of 672bp; AA066552 is the amino acid sequence deduced from the ORF.

## RESULT 5

AA42483 standard; Protein; 223 AA.

AA42483;

03-FEB-1995 (first entry)

Cynomolgus Monkey zona pellucida ZPC protein.

Contraceptive; porcine; pig; zona pellucida; mammal; lapine; canine; dog; rabbit; feline; cat; bovine; cow; human; cynomolgus monkey; ZPA; ZPB; ZPC; infertility; veterinary.

Macaca fascicularis.

US981228-A.

09-NOV-1999.

07-JUN-1995; 95US-0484596.

09-NOV-1993; 93US-0149223.

09-NOV-1992; 92US-0973341.

29-JAN-1993; 93US-0012990.

(ZONA-) ZONAGEN INC.

Podolski JS, Hsu KT, Harris JD;

WPI: 1999-633318/54.

N-PSDB; AA222740.

Mammalian zona pellucida proteins used to induce transient or permanent

infertility -

Example 12: Column 143-146; 84pp; English.

This sequence represents the cynomolgus monkey zona pellucida ZPC

protein. The invention relates to the isolation of novel nucleotide

sequences encoding zona pellucida proteins from mammalian, especially

porcine, lapine, canine, feline, bovine, human or cynomolgus monkey

sources. The zona pellucida proteins (ZPA, ZPB, and ZPC) are used in

the induction of transient or permanent infertility. At present the

method is used in veterinary applications to induce transient or

permanent infertility in porcine, lapine, canine, feline, bovine,

and cynomolgus monkeys.

Sequence 223 AA:

AAW81820 standard; Protein; 223 AA.

AAW81820;

29-JAN-1999 (first entry)

M. cynomolgus ZPC protein.

ZPC; zona pellucida; infertility; sterility; immunocontraceptive;

vacine; monkey.

Macaca cynomolgus.

Key Location/Qualifiers

1..223

US5837497-A.

17-NOV-1998.

```

X 07-JUN-1995; 95US-0484993.
X 09-NOV-1993; 93US-0149223.
X 09-NOV-1992; 92US-0973341.
X 29-JAN-1993; 93US-0012990.
X 07-JUN-1995; 95US-0484993.
X (ZONA-) ZONAGEN INC.
X Harris JD;
X WPI: 1999-023447/02.
X N-PSDB; AA64820.
X Isolated zona pellucida DNA from different mammals - used to develop
X products which can be used for vaccination to induce transient
X infertility or permanent sterility in female mammals
X Example 12; Column 145-148; 84pp; English.
X This sequence represents a Macaca cynomolgus ZPC protein isolated from
X zona pellucida. This protein can be used in a method for specifically
X inducing transient infertility or permanent sterility in a host
X animal by selective vaccination with specific zona pellucida proteins
X or immunoreceptively active fragments.
X Sequence 223 AA:
X
X Query Match 93.4%; Score 226; DB 20; Length 223;
X Best Local Similarity 90.2%; Pred. No. 5.1e-20;
X Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
X
X 1 SWFVQGPADICQCCNKDGGTSPSHSRQPHVMSQMSRSVS 41
X |||||:|||||:|||||:|||||:|||||:|||||:
X 107 swfvegpadiqccsksgdctgtpshsrqphvmsqwsrsas 147
X
X RESULT 7
X ID AAY82218 standard; Protein; 223 AA.
X AC AAY82218;
X DT 12-JUN-2000 (first entry)
X DE Monkey zona pellucida ZPC SEQ ID NO:49.
X KW zona pellucida; ZPA; ZPB; ZPC; infertility; permanent sterility;
X vaccine; immunoreceptively; contraceptive; immunostimulatory.
X OS Primate.
X PN US6027727-A.
X PD 22-FEB-2000.
X PF 09-NOV-1993; 93US-0149223.
X PR 09-NOV-1992; 92US-0973341.
X PR 29-JAN-1993; 93US-0012990.
X PA (ZONA-) ZONAGEN INC.
X PI Podolski JS, Hsu KT, Harris JD;
X DR WPI: 2000-269144/23.
X N-PSDB; AA695677.
X Inducing reproducible transient or permanent infertility in a mammal
X comprises administration of homologous and/or heterologous mammalian
X species zona pellucida proteins

```

```

XX Example 12; Column 147-148; 85pp; English.
XX The present invention describes a method for inducing reproducible
XX transient infertility in a female mammal, including humans, by selective
XX administration of homologous and/or heterologous mammalian species
XX zona pellucida (ZP) proteins or immunoreceptively active fragments.
XX Also described are: (1) a method for inducing transient infertility in a
XX mammal by administering directly an antibody directed to a ZPB or an
XX immunologically active fragment selected from feline, bovine,
XX cynomolgus monkey or human ZPB polypeptides; (2) an isolated, purified
XX recombinant feline, bovine, cynomolgus monkey or human ZPB polypeptide
XX or immunologically active fragment; and (3) a fusion protein comprising
XX a ZPB polypeptide which is conjugated with a compound selected from
XX keyhole limpet haemocyanin, muramyl dipeptide, histidine-tag, beta-gal,
XX or palmitic acid where the fusion protein remains effective to stimulate
XX production of antibodies that recognise a ZPB polypeptide. The method is
XX useful for providing transient or permanent infertility or sterility in
XX humans and animals. The present sequence represents monkey ZPC, which is
XX used in the exemplification of the present invention.
XX Sequence 223 AA:
XX
XX Query Match 93.4%; Score 226; DB 21; Length 223;
XX Best Local Similarity 90.2%; Pred. No. 5.1e-20;
XX Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 SWFVQGPADICQCCNKDGGTSPSHSRQPHVMSQMSRSVS 41
XX |||||:|||||:|||||:|||||:|||||:|||||:
XX 107 swfvegpadiqccsksgdctgtpshsrqphvmsqwsrsas 147
XX
XX RESULT 8
XX ID AAY52692 standard; Protein; 223 AA.
XX AC AAY52692;
XX DT 14-MAR-2000 (first entry)
XX DE Cynomolgus monkey oocyte zona pellucida protein ZPC.
XX KW zona pellucida; ZPA; ZPB; ZPC; purified; mammalian; glycoprotein;
XX target; immunoreceptively; vaccine; antibody; transient; infertility;
XX controllable; predictable; permanent; sterility; side effect.
XX OS Macaca cynomolgus.
XX PN US6001599-A.
XX PD 14-DEC-1999.
XX PF 02-JUN-1995; 95US-0458731.
XX PR 09-NOV-1993; 93US-0149223.
XX PR 09-NOV-1992; 92US-0973341.
XX PR 29-JAN-1993; 93US-0012990.
XX PA (ZONA-) ZONAGEN INC.
XX PI Podolski JS, Hsu KT, Harris JD;
XX DR WPI: 2000-061880/05.
XX Induced DNA encoding mammalian zona pellucida proteins useful for
XX inducing transient sterility -
XX Example 12; Columns 145-146; 84pp; English.
XX This sequence represents a mammalian zona pellucida protein, ZPC.
XX The zona pellucida is a complex matrix surrounding the mammalian
XX oocyte, formed of glycoproteins secreted by ovarian cells. The

```

CC invention relates to zona pellucida proteins ZPA, ZPB and ZPC, which  
CC are useful as targets for immunocontraceptive vaccines. Polynucleotides  
CC encoding these proteins can be used for recombinant protein production,  
CC and as probes in hybridisation methods for the isolation of  
CC polynucleotides encoding zona pellucida proteins from other mammalian  
CC species. Administration of zona pellucida proteins to a host animal  
CC results in a specific immunocontraceptive effect. Administration of  
CC purified ZPA and/or ZPB induces transient infertility in females.  
CC Fertility can be maintained in an on or off condition in a controllable/  
CC predictable fashion. Administration of ZPC induces permanent sterility.  
CC Use of a purified zona pellucida protein rather than a complex mixture  
CC of zona pellucida proteins reduces the potential for unwanted side-  
CC effects which may be harmful.

SC Sequence 223 AA:

Query Match 93.4%; Score 226; DB 21; Length 223;  
Best Local Similarity 90.2%; Pred. No. 5.1e-20;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWFPVGGPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 41  
DB 107 swfpvgpadicqcsksgdcgtpshsrqphvmsqwsrsas 147

## RESULT 9

AAV52183  
ID AAV52183 standard; Protein; 223 AA.

AC AAV52183;

DT 09-FEB-2000 (first entry)

DE Monkey zona pellucida protein C (ZPC) amino acid sequence.

KM Zona pellucida protein; ZPA; ZPB; ZPC; oocyte; ovarian cell; antibody;  
KM fusion protein; vaccine; contraceptive; fertilization; fertility;  
KM immunocontraceptive.

OS Macaca cynomolgus.

PN US5989550-A.

PD 23-NOV-1999.

PF 07-JUN-1995; 950S-0480150.

PR 09-NOV-1993; 930S-0149223.

PR 09-NOV-1992; 920S-0973341.

PR 29-JAN-1993; 930S-0012890.

PA (ZONA-) ZONAGEN INC.

PI Podolski JS, Hsu KT, Harris JD;

PT Isolated and purified mammalian zona pellucida proteins useful in  
PT methods of immunocontraception -

PS Example 12; Column 145-146; 84pp; English.

XX Sequences AAV52181-Y52183 are monkey zona pellucida proteins ZPA, ZPB,  
XX and ZPC. The invention relates to isolated and purified zona pellucida  
XX proteins from mammals, specifically canine, feline and bovine ZPA or  
XX their biologically active fragments also porcine and cynomolgus monkey  
XX proteins. The zona pellucida is a complex matrix surrounding the mammalian  
XX oocyte, formed of glycoproteins secreted by ovarian cells. The  
XX proteins, especially mentioned ZPA proteins can be used to produce a fusion  
XX protein that stimulate production of antibodies in a mammal that recognize  
XX an epitope of the mammal. The ZPA polypeptides are useful as

CC vaccines for selectively effecting transient infertility in mammals. ZP  
CC has an important role in maintaining the oocyte and in oocyte-sperm  
CC interactions and so is a target for a contraceptive agent which  
CC interferes with the fertilization process. Providing a specific  
CC polypeptide reduces the need for a complex mixture of many ZP proteins  
CC which may create unwanted and harmful side effects. The duration of  
CC transient fertility is controllable and can be maintained in a  
CC controllable and/or predictable fashion.

SC Sequence 223 AA:

Query Match 93.4%; Score 226; DB 21; Length 223;  
Best Local Similarity 90.2%; Pred. No. 5.1e-20;  
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWFPVGGPADICCCCKGDCGTPSHSRQPHVMSQMSRSVS 41  
DB 107 swfpvgpadicqcsksgdcgtpshsrqphvmsqwsrsas 147

## RESULT 10

AAV52987  
ID AAV52987 standard; Protein; 223 AA.

AC AAV52987;

DT 14-FEB-2000 (first entry)

DE Cynomolgus monkey zona pellucida ZPC protein sequence.

KM Zona pellucida; ZPA; ZPB; ZPC; infertility; sterility; vaccine;  
KM immunocontraceptive; contraception; immunisation.

OS Mammalia.

PN US5976545-A.

PD 02-NOV-1999.

PF 07-JUN-1995; 950S-0484158.

PR 09-NOV-1993; 930S-0149223.

PR 09-NOV-1992; 920S-0973341.

PR 29-JAN-1993; 930S-0012890.

PA (ZONA-) ZONAGEN INC.

PI Podolski JS, Hsu KT, Harris JD;

PT WPI: 2000-037080/03.

PT N-PSDB; AA233274.

PS New recombinant mammalian zona pellucida C proteins, for induction of  
PS female sterility -

XX Claim 14; Column 145-146; 87pp; English.

XX The present invention describes recombinant zona pellucida (ZP)  
XX proteins, specifically ZPC proteins. Also described are fusion proteins  
XX of ZPC with any of keyhole limpet haemocyanin, muramyl dipeptide,  
XX histidine tag, beta-galactosidase or palmitic acid, capable of  
XX stimulating mammalian antibodies that recognise mammalian ZPC.  
XX Recombinant ZPC proteins are used to immunise animals to induce transient  
XX infertility or permanent female sterility, in humans or other mammals.  
XX AA233243 to AA233254, AA233270 to AA233274 and AA233285 encode mammalian  
XX ZP proteins given in AAV52971 to AAV52988 from the present invention.  
XX AA233255 to AA233269 and AA233275 to AA233284 represent oligonucleotides  
XX used in the exemplification of the present invention.

SC Sequence 223 AA:

us-09-252-828a-2.rag

ID AAR53498 Standard, Protein, 423 mg  
 XX  
 AC AAR53498;  
 XX

DI 02-NOV-1994 (first entry)  
XX  
DE Marmoset ZP3.  
XX

XX

Callitriche juccatus.  
XX  
05

XX

PD 11-MAY-1954.

XX

XX	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
XX	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

PR 02-NOV-1992; 32EF 0  
XX  
PA (AKZII ) AKZO NOBEL, NY

AA Aitken RJ, Koothan T;  
PI

DR WPI; 1994-167461/20

XX

immuno-contraceptive vaccines and, in diagnosis, also test systems for detecting development

XX  
PC  
Claim 1: Page 15: 27nn: English

XX  
CC The marmoset ZP3 gene was isolated by PCR from total mammoset RNA

(the sperm receptor) is useful as an immun contraceptive and as

CC effects of vaccination, for imaging follicles, to detect ZP3-re-

CC autoantibodies against ZP3.

SQ Sequence 4.24 AA;

Query Match 79.3%; Score 192; DB 15; Length 424;

Matches	32;	Conservative	4;	Mismatches	3;	Indels	0;	Gaps	0
---------	-----	--------------	----	------------	----	--------	----	------	---

```

QY      1  SNFPPVGGPADICQCCCKNGCCGPIPSHSRRQPHVMSQMSRS 39
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      308  swfpevgpadicqccskgdcgfpsharrqphvslsgss 346

RESULT  13
AAK55199
ID      AAK55199 standard; Protein; 426 AA.
XX
AC      AAK55199;

```

DT 31-JAN-1995 (first entry)

XX canine zona pellucida; ZPC; immunoneutralization  
XX Dog; canine; zona pellucida; ZPC; immunoneutralization  
XX

XX.	key	Location/Qualifiers
FH	Protein	1..426
FT		

AA WO9411019-A  
PN



•

CC This sequence represents a canine ZPC protein isolated from zona  
CC pellucida. This protein can be used in a method for specifically  
CC inducing transient infertility or permanent sterility in a host  
CC animal by selective vaccination with specific zona pellucida proteins  
CC or immunocontraceptively active fragments.

Sequence 426 AA;

